

STIC-Biotech/ChemLib

121194

From: Rao, Manjunath N.  
Sent: Wednesday, May 05, 2004 10:05 AM  
To: STIC-Biotech/ChemLib  
Subject: Sequence search request for 10/088,676

From: Manjunath N. Rao  
Art Unit 1652, Room 3B81  
Mail Box in Room 3C70  
Phone: 272-0939

Date: 5-5-04

Please search the following as soon as possible for application with serial number

**10/088,676**

1. SEQ ID NO: 1 against all commercial nucleic acid databases, issued patents/published applications database and pending application database. Please provide a print of all results
2. SEQ ID NO: 2, against all commercial amino acid databases, issued patents/published applications database and pending application database. Please provide a print of all results.

If you have any questions please call me at the above phone number.

Thanks

Manjunath N. Rao  
Art Unit 1652, Room 3B81,  
Mail Box in Room 3C70,  
Remsen Building, USPTO

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: 5/14  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

400, Dulany St.  
Alexandria, VA.  
Phone: 571-272-0939

RECEIVED  
MAY - 5 2004  
(STIC)

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: \_\_\_\_\_  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

Page 4

Query	Best Local Similarity	99.8%	Score 1775.8	DB 4	Length 2060
Matches 1777	Conservative	0	Mismatches 2	Indels 0	Gaps 0
Query	1	ATGAGGATGCTTGGTGGCTTCCCTGAGACCAATGCTCCGACGAACTCCGCGCCCGCGCG	60		
Db	26	ATGAGGATGCTTGGTGGCTTCCCTGAGACCAATGCTCCGACGAACTCCGCGCCCGCG	85		
Query	61	TGCTTACGCCCCGGGGGCTCTCTACTTGGCTCTGTGTCCATCTCTCCCTTCTCCGAG	120		
Db	86	TGCTTACGCCCCGGGGGCTCTCTACTTGGCTCTGTGTCCATCTCTCCCTTCTCCGAG	145		
Query	121	GCTGAGACAGAGAGACCTTGGCTGCTGTACAGAGCTGAGAGTTGAGAGAAAGACCTG	180		
Db	146	GCTGAGACAGAGAGACCTTGGCTGCTGTACAGAGCTGAGAGTTGAGAGAAAGACCTG	205		
Query	181	ATTCTACTTGTATGTAGAGACCAAGAACCACTCAGAGACATCATATGAACTTCTCTCT	240		
Db	206	ATTCTACTTGTATGTAGAGACCAAGAACCACTCAGAGACATCATATGAACTTCTCTCT	265		
Query	241	CTGCACTGAGATCCGTCATCATTTATATATGGCTGGCTCGATTTCTTAACTCCAGCGC	300		
Db	266	CTGCACTGAGATCCGTCATCATTTATATATGGCTGGCTCGATTTCTTAACTCCAGCGC	325		
Query	301	TTGGTACACCTGAGCGGGGACCTTCCGCGCGCTTTCTGCGCTTCCGGGGCAAAAGAC	360		
Db	326	TTGGTACACCTGAGCGGGGACCTTCCGCGCGCTTTCTGCGCTTCCGGGGCAAAAGAC	385		
Query	361	GACTTCTGTCAGATTCAGAACCTTAGAGAACCCGGCGAAAGCCGGCGGGCCCGGCGCG	420		
Db	386	GACTTCTGTCAGATTCAGAACCTTAGAGAACCCGGCGAAAGCCGGCGGGCCCGGCGCG	445		
Query	421	GATTAATATCTGAAACATATGAGAGATGACATTTGTGAGAGTATGTGTGCTTAGATAA	480		
Db	446	GATTAATATCTGAAACATATGAGAGATGACATTTGTGAGAGTATGTGTGCTTAGATAA	505		
Query	481	CAGAAAGGCTGCAAGATTGCCAGACCCCTGATGTATGCTGTGCTCCAAAGAGAGAG	540		
Db	506	CAGAAAGGCTGCAAGATTGCCAGACCCCTGATGTATGCTGTGCTCCAAAGAGAGAG	565		
Query	541	GCAGCTCAGATGATCTGGTCTTTTAAAGAGCAATTCCTCAATATTTCAGTAATCTC	600		
Db	566	GCAGCTCAGATGATCTGGTCTTTTAAAGAGCAATTCCTCAATATTTCAGTAATCTC	625		

QY 601 ATATTAACAGCCGAGTCTCTAGACAACTTTATTAATCTTGATGCTCTGACCTCCAC 660  
 DB 626 ATTTTAACAGCCGAGTCTCTAGACAACTTTATTAATCTTGATGCTCTGACCTCCAC 685  
 QY 661 CTGATATTTGCTCTAAATGACCTGCGTCTGATCCCAATTAATCTCTGAAACAGTTCTAGT 720  
 DB 686 CTGATATTTGCTCTAAATGACCTGCGTCTGATCCCAATTAATCTCTGAAACAGTTCTAGT 745  
 QY 721 GCCCTGAGTCTGTTGAATGACAGCCGACGAAATTAATTAATCTTTCTGGAATCTGGGT 780  
 DB 746 GCCCTGAGTCTGTTGAATGACAGCCGACGAAATTAATTAATCTTTCTGGAATCTGGGT 805  
 QY 781 AATGAGCCAAATTAATCTGATGACCAATGATGCGCGGAGTAAATGGACGCAATGGGA 840  
 DB 806 AATGAGCCAAATTAATCTGATGACCAATGATGCGCGGAGTAAATGGACGCAATGGGA 865  
 QY 841 AAGGATTAATCACTGATGAAAGACCTGTTGACAGCCCATCCGATTTATTCAGAGCCAGC 900  
 DB 866 AAGGATTAATCACTGATGAAAGACCTGTTGACAGCCCATCCGATTTATTCAGAGCCAGC 925  
 QY 901 TTATATGCGCCCTAATATGCGCGCGCGAGGAGAAATGATGCGCCCTCTAGATGATTC 960  
 DB 926 TTATATGCGCCCTAATATGCGCGCGCGAGGAGAAATGATGCGCCCTCTAGATGATTC 985  
 QY 961 ATGAGAGTGGCAGAAAGTACAGTAGATGAGTTCCTGCAACATTTGCTACATGATGAGC 1020  
 DB 986 ATGAGAGTGGCAGAAAGTACAGTAGATGAGTTCCTGCAACATTTGCTACATGATGAGC 1045  
 QY 1021 CCGGTGCTCAAGGTATGATGATCTCTGAAATCTCGCTGTTGACACACTCTCTGACAG 1080  
 DB 1046 CCGGTGCTCAAGGTATGATGATCTCTGAAATCTCGCTGTTGACACACTCTCTGACAG 1105  
 QY 1081 ATTAGGAAATTCAGAAAGTGTGTTAATACATACATCCGAGAAAGATTTGGCTTGA 1140  
 DB 1106 ATTAGGAAATTCAGAAAGTGTGTTAATACATACATCCGAGAAAGATTTGGCTTGA 1165  
 QY 1141 GGTGTGTGACCACTGACCTGAGGAGCAAAACATTAATCCGATTCCTATGCTGACAGA 1200  
 DB 1166 GGTGTGTGACCACTGACCTGAGGAGCAAAACATTAATCCGATTCCTATGCTGACAGA 1225  
 QY 1201 TTTTATGTTGACCACTTTAGAAATGCTGGGCAATCGAGGATGATGCTGATAGG 1260  
 DB 1226 TTTTATGTTGACCACTTTAGAAATGCTGGGCAATCGAGGATGATGCTGATAGG 1285  
 QY 1261 CACTCATTTTGTGACCATGATATCAATCACTCGTGACACAGATTTTAACCATTAACA 1320  
 DB 1286 CACTCATTTTGTGACCATGATATCAATCACTCGTGACACAGATTTTAACCATTAACA 1345  
 QY 1321 GACTATGCTCTCTCTCTCTCAAAAGGCTGATGCGGCCCAAAAGTCTTGCTGTGAT 1380  
 DB 1346 GACTATGCTCTCTCTCTCTCAAAAGGCTGATGCGGCCCAAAAGTCTTGCTGTGAT 1405  
 QY 1381 GTGGCTGGGCTCTCAGGAGGACGAGGCTGCGGAGTATCCGGGACAACTTAAGATT 1440  
 DB 1406 GTGGCTGGGCTCTCAGGAGGACGAGGCTGCGGAGTATCCGGGACAACTTAAGATT 1465  
 QY 1441 TATGCTACTGCAAAACCAACCAACCAACCACTTCTGCTGCTTCACTTTT 1500  
 DB 1466 TATGCTACTGCAAAACCAACCAACCAACCACTTCTGCTGCTTCACTTTT 1525  
 QY 1501 ATCATCAATCTGATGATCAAGAAAGAAATCAAGCTGCTGGGATCTCTAGAGACAG 1560  
 DB 1526 ATCATCAATCTGATGATCAAGAAAGAAATCAAGCTGCTGGGATCTCTAGAGACAG 1585  
 QY 1561 CTGTTTACAGTACTGCTGACAGCCCTAATGGGACGAGGCTTAAATGCCAATGATG 1620  
 DB 1586 CTGTTTACAGTACTGCTGACAGCCCTAATGGGACGAGGCTTAAATGCCAATGATG 1645  
 QY 1621 CAACGATATGGGACCCCTTATGATGATGATGATGATGATGATGATGATGATGATG 1680  
 DB 1646 CAACGATATGGGACCCCTTATGATGATGATGATGATGATGATGATGATGATGATG 1705

QY 1681 GCCCCTTGGGCGCGCGGACAT ACATGGGCTTTATG 1740  
 DB 1706 CGCCCTTGGGCGCGCGGACATGCTCATCCCTCCAGTACATAGGCTTTTGTG 1765  
 QY 1741 GTCAAGATGTCATGCTTTGGCGCGCTAGCGATTA 1779  
 DB 1766 GTCAAGATGTCATGCTTTGGCGCGCTAGCGATTA 1804

RESULT 4  
 AAH22671  
 ID AAH22671 standard; DNA; 2636 BP.  
 XX  
 AC AAH22671;  
 DT 07-SEP-2001 (first entry)  
 DE Heparanase-like protein Hpa2 splice variant #1 encoding DNA.  
 KW Heparanase; splice variant; homologue; heparanase-like protein; Hpa2;  
 KW cytosolic; neuroprotective; cerebroprotective; immunosuppressive;  
 KW antidiabetic; antiarteriosclerotic; vulnery; ds.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT 5'UTR 1..600  
 FT CDS /\*tag= a  
 FT 601..2379 /\*tag= b  
 FT /transl\_except= (pos: 1309..1311, aa: Xaa)  
 FT /note= "Xaa = unknown"  
 FT 3'UTR 2377..2636  
 FT /\*tag= c

MO200146392-A2.  
 28-JUN-2001.  
 21-DEC-2000; 2000NO-GB004963.  
 22-DEC-1999; 99GB-00030392.  
 07-APR-2000; 2000GB-00008713.  
 (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
 PI McKenzie EA, Stamps AC, Terrett JA, Tyson KL;  
 WI; 2001-418056/44.  
 DR P-PSDB; AAB85215.  
 DR Novel homologs of heparanase, present in three splice variants, useful  
 PT for identifying agents that modulate heparanase, useful in the treatment  
 PT and/or prophylaxis of abnormal levels of heparanase.  
 XX  
 PS Claim 12; Fig 1; 97pp; English.

The invention provides a homologue to heparanase which is present in  
 CC three splice variants. The heparanase homologue polypeptide is useful in  
 CC the treatment of a human or non-human animal or for use in diagnosis.  
 CC Vectors comprising the heparanase homologue polynucleotides are useful in  
 CC the transformation or transfection of a prokaryotic or eukaryotic host.  
 CC The modulators of the polypeptide are useful in the manufacture of a  
 CC medicament for the treatment and/or prophylaxis of a condition/disease  
 CC associated with abnormal levels of the heparanase homologue, including  
 CC cancer, central nervous system (CNS) and neurodegenerative diseases,  
 CC cardiovascular diseases such as restenosis following angioplasty and  
 CC atherosclerosis, autoimmune diseases, psoriasis, lupus erythematosus,  
 CC allografts, inflammatory diseases, arthritis, vascular restenosis, tumour  
 CC growth and progression, asthma, Alzheimer's disease, diabetic  
 CC retinopathy, wound healing and inflammation. The polypeptide is also  
 CC useful in diagnosis and research. The present sequence represents the



sequence encoding it. Heparanase-2 is a member of the endoglyucuronidase family of polypeptides and it degrades heparan sulphate proteoglycans HSPGs (biglycan) macromolecules of cell surfaces, basement membranes and the extracellular matrix). HSPGs support the vascular endothelium and stabilises the structure of the capillary wall. Heparanases may be associated with neovascularisation and metastasis related to malignant tumour formation. Heparanase-2 polynucleotides and proteins are useful as vaccines for inducing an immunological response against autoimmune disorders, blood coagulation disorders, cancer, diabetes, ichthemia, sepsis, stroke, cardiovascular diseases, or thrombosis, as well as in diagnosing (the susceptibility of a subject to) these diseases.

Heparanase-2 fragments may be used as immunogens to produce antibodies immunospecific to the polypeptides, and to identify membrane bound soluble receptors, agonists or antagonists that compete with the binding of the polypeptide to the receptor. An antibody specific for heparanase-2 can be used in the diagnosis of the above diseases and in isolating or identifying clones expressing heparanase-2. The present sequence represents heparanase-2. Three regions of heparanase-2 with high immunogenicity (immunogenic epitopes) can be used to raise antibodies against heparanase-2.

SQ Sequence 592 AA;

Query Match	100.0%;	Score 3088;	DB 4;	Length 592;
Best Local Similarity	100.0%;	Pred. No. 1.3e-304;		
Matches 592; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

Qy	1	MRVYCAPPEAMPSSNSRPPACLAAGLYALILHHSLSAQADRRPLPVDRAAGKEKTL	60
Dd	1	MRVYCAPPEAMPSSNSRPPACLAGLYALILHHSLSAQADRRPLPVDRAAGKEKTL	60
Qy	61	ILLDVSTKNRVRTNENFSLSLQDPSIHHGMDLFSRRLLTTLARGLSPAFLEGKRT	120
Dd	61	ILLDVSTKNRVRTNENFSLSLQDPSIHHGMDLFSRRLLTTLARGLSPAFLEGKRT	120
Qy	121	DFLOFONLRNPAPKSRGGPGPDYILKRYEDDIYRSVDALDXKXGCKIAQHPDVMYLQREK	180
Dd	121	DFLOFONLRNPAPKSRGGPGPDYILKRYEDDIYRSVDALDXKXGCKIAQHPDVMYLQREK	180
Qy	181	AAQWHLTLKQESNTYSNLILTAASLDKLYNADCSGLHFLPANAIRBNPNNSWNSS	240
Dd	181	AAQWHLTLKQESNTYSNLILTAASLDKLYNADCSGLHFLPANAIRBNPNNSWNSS	240
Qy	241	ALSLTKYSASKKNI SWELGNEBNRYTHGKRAVNSQLGXDIOLKSLQPIRYSRAS	300
Dd	241	ALSLTKYSASKKNI SWELGNEBNRYTHGKRAVNSQLGXDIOLKSLQPIRYSRAS	300
Qy	301	LYGNBGRPKKNVIALIDGFMKVASGTDAVTWQHCYIDGRVYKXMDPLKTRLLDPTLSDQ	360
Dd	301	LYGNBGRPKKNVIALIDGFMKVASGTDAVTWQHCYIDGRVYKXMDPLKTRLLDPTLSDQ	360
Qy	361	IRKI QKXVNYTPEKKI TWELGVTTSAGGNNLSDSYAGFPLMTLIGLANOGIDVYIR	420
Dd	361	IRKI QKXVNYTPEKKI TWELGVTTSAGGNNLSDSYAGFPLMTLIGLANOGIDVYIR	420
Qy	421	HSFPDHGNIHVDQNFPLPDYMLSLYKRLGPKYLAVHVAQORPKRPGVIRIDKLRI	480
Dd	421	HSFPDHGNIHVDQNFPLPDYMLSLYKRLGPKYLAVHVAQORPKRPGVIRIDKLRI	480
Qy	481	YAHCTNHNHNYYVGSITLFIINLHRSKKIKLAGTLRDLVHGYLLQPYGQEGLSKSV	540
Dd	481	YAHCTNHNHNYYVGSITLFIINLHRSKKIKLAGTLRDLVHGYLLQPYGQEGLSKSV	540
Qy	541	QLNQOPLVMDDGLPELXKPRPLAAGTLYIPVWTMGFYVNVNVALAQRYS	600
Dd	541	QLNQOPLVMDDGLPELXKPRPLAAGTLYIPVWTMGFYVNVNVALAQRYS	600

XX	20-APR-2001 (first entry)
DT	
XX	
DE	Human heparanase, hnhp1, 1
XX	

KM Hepatocarcinoma; wound healing; angiogenesis; restenosis; Scrape;  
KM atherosclerosis; inflammation; pulmonary disease; Alzheimer's disease  
KM atherosclerotic disease; Creutzfeldt-Jakob disease; viral infection;  
KM gene therapy; human.

OS	Homo sapiens.
XX	
PN	WO200100643-A2.

PD	04-JAN-2001.
XX	
PF	19-JUN-2000; 2000WO-IL000358

PR 25-JUN-1999; 99US-0140801P.  
XX  
PA (INSI-) INSIGHT STRATEGY & MARKETING LTD

PT New polynucleotides and polypeptides that are distantly homologous to  
PT heparanase, useful in wound healing, as well as in gene therapy protocols  
PT for angiogenesis, restenosis, atherosclerosis, or inflammation.

PS Claim 10; Fig 1; 67pp; English

CC This sequence represents a heparanase of the invention. The heparanase  
CC DNA and protein sequences are useful in wound healing, angiogenesis,  
CC retinosis, atherosclerosis, inflammation, pulmonary disease,  
CC neurodegenerative diseases (such as scrapie, Alzheimer's disease, and  
CC Creutzfeldt-Jakob disease) or viral infections. The heparanase coding  
CC sequence is particularly useful in gene therapy

**SQ Sequence 592 AA;**

Query Match	99.7%;	Score 3078;	DB 4;	Length 552;
Best Local Similarity	99.7%;	Pred. No. 1.3e-303;		
Matches 590;	Conservative 1;	Mismatches 1;	Indels 0;	Gaps 0;

Oy	1	MVLTCAFPMAMSSSRPPACIACGALYALMLHLSTSSOAGRRRLPVDRAAGLKEKTL	60
Db	1	MVLTCAFPMAMSSSRPPACIACGALYALMLHLSTSSOAGRRRLPVDRAAGLKEKTL	60
Oy	61	ILLDVSTKPVATVNEFLSLQDPSIIHOGMDFLSSKSLVTLAAGLSAPFLRFQGGKT	120
Db	61	ILLDVSTKPVATVNEFLSLQDPSIIHOGMDFLSSKSLVTLAAGLSAPFLRFQGGKT	120
Oy	121	DLFQFQNLNPNPAKSGGSGGPDYYLKNTEDDIVRSVDALDKQKCGKTAQHPDVMVLQREK	180
Db	121	DLFQFQNLNPNPAKSGGSGGPDYYLKNTEDDIVRSVDALDKQKCGKTAQHPDVMVLQREK	180
Oy	181	AAQMHVLLKEQFSMTYINLLITARSIDLKYNFADCGGHLITPLNLAHPNPNMNSSS	240
Db	181	AAQMHVLLKEQFSMTYINLLITARSIDLKYNFADCGGHLITPLNLAHPNPNMNSSS	240
Oy	241	ALSLTKYSASKKCNISMEIAGNEPNNYRTMGRANVSGQLCKDYIOLKSLQPIRISRAS	300
Db	241	ALSLTKYSASKKCNISMEIAGNEPNNYRTMGRANVSGQLCKDYIOLKSLQPIRISRAS	300
Oy	301	LYGPNTGRPRKAVIALLDGFMKVASGYDAVTMOHCITIDGRVYKVMDFELTRLIDTLISQ	360
Db	301	LYGPNTGRPRKAVIALLDGFMKVASGYDAVTMOHCITIDGRVYKVMDFELTRLIDTLISQ	360
Oy	361	IRIKGVNNTYRGGKKIMLEGVYTSAGCTNNLSDSYAAGFIMLNTGLMANGCIDVVR	420
Db	361	IRIKGVNNTYRGGKKIMLEGVYTSAGCTNNLSDSYAAGFIMLNTGLMANGCIDVVR	420





CC family of polypeptides and it degrades heparan sulphate proteoglycans  
CC HSPGs (ubiquitous macromolecules of cell surfaces, basement membranes and  
CC the extracellular matrix). HSPGs support the vascular endothelium and  
CC stabilise the structure of the capillary wall. Heparanases may be  
CC associated with neovascularisation and metastasis related to malignant  
CC tumour formation. Heparanase-2 polynucleotides and proteins are useful as  
CC vaccines for inducing an immunological response against autoimmune  
CC disorders, blood coagulation disorders, cancer, diabetes, ischaemia,  
CC sepsis, stroke, cardiovascular diseases, or thrombosis, as well as in  
CC diagnosing (the susceptibility of a subject to) these diseases.  
CC Heparanase-2 fragments may be used as immunogens to produce antibodies  
CC heparanase-specific to the polypeptides, and to identify membrane bound  
CC soluble receptors, agonists or antagonists that compete with the binding  
CC of the polypeptide to the receptors. An antibody specific for heparanase-  
CC 2 can be used in the diagnosis of the above diseases and in isolating or  
CC identifying clones expressing heparanase-2. The present sequence  
CC represents cDNA encoding heparanase-2

Sequence 1779 BP; 443 A; 477 C; 437 G; 422 T; 0 U; 0 Other;

Query Match	Score	DB	Length
100.0%	1779	4	1779

Best Local Similarity 100.0%; Pred. No. 0;  
Matched 1779; Corresponding A: Matched

Matches 1779; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	ATGAGGGTGTGTTGTGCTTCCCTGAAGGCAATGCCCTCAGAACTCCGGCCCCCGG	60
Dp	1 ATGAGGGTGTGTTGTGCTTCCCTGAAGGCAATGCCCTCAGAACTCCGGCCCCCGG	60
Qy	61 TGCCCTAGCCCCGGGGGGCTCTCTAATTGAGCTGTGTGCTCAATCTCCCTTCTCCAG	120
Dp	61 TGCCCTAGCCCCGGGGGGCTCTCTAATTGAGCTGTGTGCTCAATCTCCCTTCTCCAG	120
Qy	121 GCTGGAGACAGAGAACCCCTTGCTGTGAGCAGAAGCTGAGATTGAAAGGAAAATAACCTG	180
Dp	121 GCTGGAGACAGAGAACCCCTTGCTGTGAGCAGAAGCTGAGATTGAAAGGAAAATAACCTG	180
Qy	181 ATTCTACTGATGTGAGCAACAAAGAACCAAGTCAGAGACGTCAATTGAAATCTTCTCT	240
Dp	181 ATTCTACTGATGTGAGCAACAAAGAACCAAGTCAGAGACGTCAATTGAAATCTTCTCT	240
Qy	241 CTGAGAGTGAATCCGTCACATATATATATGCTGGCTCAATTCCTAAAGTCCAAAGCG	300
Dp	241 CTGAGAGTGAATCCGTCACATATATATATGCTGGCTCAATTCCTAAAGTCCAAAGCG	300
Qy	301 TTGGTGAACCTGGCCCGGGAGCTTTTCGCCGCTTTGAGCCTTCGGGGGCAAAAGAAC	360
Dp	301 TTGGTGAACCTGGCCCGGGAGCTTTTCGCCGCTTTGAGCCTTCGGGGGCAAAAGAAC	360
Qy	361 GACTTCTGTGAGTTCCAGAACCTGAGGAACCCGCGAAGAACCCGCGGGGCGCGG	420
Dp	361 GACTTCTGTGAGTTCCAGAACCTGAGGAACCCGCGAAGAACCCGCGGGGCGCGG	420
Qy	421 GATTACTATCTCAAAAACATGAGATGACATGTGTCAGAGTATGTTGCTTGAATPAA	480
Dp	421 GATTACTATCTCAAAAACATGAGATGACATGTGTCAGAGTATGTTGCTTGAATPAA	480
Qy	481 CAGAAAGGCTGCAAGATTGCCAGACACCTGTGATGATGCTGCTCCAAAGGAGAG	540
Dp	481 CAGAAAGGCTGCAAGATTGCCAGACACCTGTGATGATGCTGCTCCAAAGGAGAG	540
Qy	541 GCAGCTCAGATGATCTGGTCTCTGAAAAGAGCAATCTCCAAATCACTTACAGTAATCTC	600
Dp	541 GCAGCTCAGATGATCTGGTCTCTGAAAAGAGCAATCTCCAAATCACTTACAGTAATCTC	600
Qy	601 AATTTAAACAGCAGAGTCTCTAGCAAACTTTATTAATTGTGATGCTCTGACCTCAC	660
Dp	601 AATTTAAACAGCAGAGTCTCTAGCAAACTTTATTAATTGTGATGCTCTGACCTCAC	660
Qy	661 CTGATATTTGCTCTAAATGACATGGGTGTATATCCAAATATCTCTGGAAAGATTTAGT	720
Dp	661 CTGATATTTGCTCTAAATGACATGGGTGTATATCCAAATATCTCTGGAAAGATTTAGT	720
Qy	721 GCCCTAGTCTGTGAAGTACAGCGCAGCAAAAAGTACAACTTTCTTGGGAATGGGT	780

Db	721	GCCTGAGTCTGTGAACTACAGCGCCAGCAAAAGTACAACTTTCTTGGAACTGGT	780
OY	781	AATAGCCAAATTAATACTATCGGACCATGCTGCGGCGAGTAAATGSCAGCCAGTTGGGA	840
Db	781	AATAGGCCAAATTAATACTATCGGACCATGATGCGCGGAGATAAATGCGACGACAGTTGGGA	840
OY	841	AAGATTACATCCAGCTGGAAGAGCGCTGTGCAAGCCATCCGGATTATATCCAGAGCCAGC	900
Db	841	AAGATTACATCCAGCTGGAAGAGCGCTGTGCAAGCCATCCGGATTATATCCAGAGCCAGC	900
OY	901	TTATATGCGCCTTAATAATTGGCGCGCCAGAGAAATGATCATCGCCCTCTAGATGGAATTC	960
Db	901	TTATATGCGCCTTAATAATTGGCGCGCCAGAGAAATGATCATCGCCCTCTAGATGGAATTC	960
OY	961	ATGAAAGGTGACAGAAAGTACAGTATGATGCAATTCCTGAGCAATTCCTGATGATGAGC	1020
Db	961	ATGAAAGGTGACAGAAAGTACAGTATGATGCAATTCCTGAGCAATTCCTGATGATGAGC	1020
OY	1021	CGGCTGATCAAGGTGATGGAATTTCTTGAAAATCGCCTGTTAGACACACTCTTCGACCG	1080
Db	1021	CGGCTGATCAAGGTGATGGAATTTCTTGAAAATCGCCTGTTAGACACACTCTTCGACCG	1080
OY	1081	ATTAGAAATATTCAGAAAGTGGTTAATACATACCTCCAGAGAAAGAAATTTGGCTTGA	1140
Db	1081	ATTAGAAATATTCAGAAAGTGGTTAATACATACCTCCAGAGAAAGAAATTTGGCTTGA	1140
OY	1141	GGTGTGTGACCACTCAGCTGAGGAGCAAAACATCTATCCGATTCCTATGCTGACAGA	1200
Db	1141	GGTGTGTGACCACTCAGCTGAGGAGCAAAACATCTATCCGATTCCTATGCTGACAGA	1200
OY	1201	TTCTTATGTTGAAACACTTTAGAGATGCTGGCCATCAGGGCATTTGATGCTGATACGG	1260
Db	1201	TTCTTATGTTGAAACACTTTAGAGATGCTGGCCATCAGGGCATTTGATGCTGATACGG	1260
OY	1261	CACGCAATTTTGGACATGATTAACATCACTCGTGGAGACGAAATTTTAAACCATTAACA	1320
Db	1261	CACGCAATTTTGGACATGATTAACATCACTCGTGGAGACGAAATTTTAAACCATTAACA	1320
OY	1321	GACTACTGGCTCTCTCTCTCTTACAAAGCGCTGATCGGCCCCAAAGTCTTGAGCTGTGAT	1380
Db	1321	GACTACTGGCTCTCTCTCTCTTACAAAGCGCTGATCGGCCCCAAAGTCTTGAGCTGTGAT	1380
OY	1381	GTGGCTGGGCTCCAGCGGAAGCCAGCGCTGAGCGAGTATCCGGGACAAACTAAAGAT	1440
Db	1381	GTGGCTGGGCTCCAGCGGAAGCCAGCGCTGAGCGAGTATCCGGGACAAACTAAAGAT	1440
OY	1441	TATGCTCATGCAAAACCAACAACAACAACAACAATCACTGCTGGGTCCATTAACCTTTT	1500
Db	1441	TATGCTCATGCAAAACCAACAACAACAACAACAACAATCACTGCTGGGTCCATTAACCTTTT	1500
OY	1501	ATCATCAACTGCACTGATCAAGAAAGAAATCAAGCTGGCTGGGACTCTCAGAGACAAG	1560
Db	1501	ATCATCAACTGCACTGATCAAGAAAGAAATCAAGCTGGCTGGGACTCTCAGAGACAAG	1560
OY	1561	CTGCTTCAACGATACCTGCTGCAAGCCTTATGAGGAGAGGAGCTTAAAGTCCAAGTCAAG	1620
Db	1561	CTGCTTCAACGATACCTGCTGCAAGCCTTATGAGGAGAGGAGCTTAAAGTCCAAGTCAAG	1620
OY	1621	CACTGATGAGCCAGCCTTATGATGATGATGAGACAGAGGAACTCCCAAGATTTGAAGCC	1680
Db	1621	CACTGATGAGCCAGCCTTATGATGATGATGAGACAGAGGAACTCCCAAGATTTGAAGCC	1680
OY	1681	CGCCCCCTTGGGCGCGCGGACATGATGATGATGATGATGATGATGATGATGATGATGATG	1740
Db	1681	CGCCCCCTTGGGCGCGCGGACATGATGATGATGATGATGATGATGATGATGATGATGATG	1740
OY	1741	GTCAAGAAATGCAATGCTTTGGCTGCGGCTACCGATTA	1779
Db	1741	GTCAAGAAATGCAATGCTTTGGCTGCGGCTACCGATTA	1779

## RESULT 2







QY 601 AATATTAAACAGCAGGCTCTCTAGACAACTTATTAATTGTGATGCTCTGAGCTCCAC 660  
 DB 626 AATATTAAACAGCAGGCTCTCTAGACAACTTATTAATTGTGATGCTCTGAGCTCCAC 685  
 QY 661 CTGATATTTTGTCTTAATGACATGCGCTGCTAATCCCAATTAATCTCTGGAACGTTCTAGT 720  
 DB 686 CTGATATTTTGTCTTAATGACATGCGCTGCTAATCCCAATTAATCTCTGGAACGTTCTAGT 745  
 QY 721 GCGCTGAGTCTGTGAAGTACAGCGCCAGCAAAAATGACAACTTCTTGGAACCTGAGT 780  
 DB 746 GCGCTGAGTCTGTGAAGTACAGCGCCAGCAAAAATGACAACTTCTTGGAACCTGAGT 805  
 QY 781 AATGAGCCAAATTAATGATGAGGACCATGATGCGCGGCGAGTAAATGGCAGCCAGTTGGGA 840  
 DB 806 AATGAGCCAAATTAATGATGAGGACCATGATGCGCGGCGAGTAAATGGCAGCCAGTTGGGA 865  
 QY 841 AAGGATTTACATCGAGCTGAAGAGCCTGTGAGCAGCCCATCGGATTTATTCAGAGCCAGC 900  
 DB 866 AAGGATTTACATCGAGCTGAAGAGCCTGTGAGCAGCCCATCGGATTTATTCAGAGCCAGC 925  
 QY 901 TTATATGCCCCCTAATAATTGGGCGGCGGAGAGAAATGTCATGCGCCTCTAGATGATTC 960  
 DB 926 TTATATGCCCCCTAATAATTGGGCGGCGGAGAGAAATGTCATGCGCCTCTAGATGATTC 985  
 QY 961 ATGAGGTGCGAAGAGTACAGTAAATGAGTAACTGAGCAATGCTGCAATGATGATGAGC 1020  
 DB 986 ATGAGGTGCGAAGAGTACAGTAAATGAGTAACTGAGCAATGCTGCAATGATGATGAGC 1045  
 QY 1021 CCGGCTGCTCAAGGTGATGCACTTCTGAAAACCTGCGCTGTGAGACACACTCTCTGACAG 1080  
 DB 1046 CCGGCTGCTCAAGGTGATGCACTTCTGAAAACCTGCGCTGTGAGACACACTCTCTGACAG 1105  
 QY 1081 AATGAGAAATTCAGAAAGTGTGTTAATACATCACTCCAGAAAAGAAATTTGGCTTGA 1140  
 DB 1106 AATGAGAAATTCAGAAAGTGTGTTAATACATCACTCCAGAAAAGAAATTTGGCTTGA 1165  
 QY 1141 GGTGTGTGACACCTCAGCTGAGGAGCAAAACATCTATCCGATTCATATGCTGAGAGA 1200  
 DB 1166 GGTGTGTGACACCTCAGCTGAGGAGCAAAACATCTATCCGATTCATATGCTGAGAGA 1225  
 QY 1201 TTCTTATGTTGAACACTTTAGGAATGCTGGCAATCAGGCGATTTGATGCTGATACGG 1260  
 DB 1226 TTCTTATGTTGAACACTTTAGGAATGCTGGCAATCAGGCGATTTGATGCTGATACGG 1285  
 QY 1261 CACTCAATTTTTCAGCATGATCAATCACTGCTGAGACAGAAATTTAACCCATTACCA 1320  
 DB 1286 CACTCAATTTTTCAGCATGATCAATCACTGCTGAGACAGAAATTTAACCCATTACCA 1345  
 QY 1321 GACTACTGCTCTCTCTCTCTAACAAGGCTGATGCGCCCAAGTCTTGAGCTGTGACAT 1380  
 DB 1346 GACTACTGCTCTCTCTCTCTAACAAGGCTGATGCGCCCAAGTCTTGAGCTGTGACAT 1405  
 QY 1381 GTGCTGTGGCTCCAGCGGAGCCAGCGCTGCGCGATGATCCGGGAGCAAACTTAAAGATT 1440  
 DB 1406 GTGCTGTGGCTCCAGCGGAGCCAGCGCTGCGCGATGATCCGGGAGCAAACTTAAAGATT 1465  
 QY 1441 TATGCTCACTGACAAACACACACAACTAATGCTGTGGATTCATTACATCTTTT 1500  
 DB 1466 TATGCTCACTGACAAACACACACAACTAATGCTGTGGATTCATTACATCTTTT 1525  
 QY 1501 ATCATCAACTGTCATGATCAAGAAAGAAATCAAGTGTGGCTGAGATCTCAAGACAG 1560  
 DB 1526 ATCATCAACTGTCATGATCAAGAAAGAAATCAAGTGTGGCTGAGATCTCAAGACAG 1585  
 QY 1561 CTGCTTACACAGTACTGCTGAGCCCTAATGGGAGAGGAGGCTTAAAGTCCAAAGTCAAGT 1620  
 DB 1586 CTGCTTACACAGTACTGCTGAGCCCTAATGGGAGAGGAGGCTTAAAGTCCAAAGTCAAGT 1645  
 QY 1621 CAATGTAATGGCCGCTTAAATGATGATGAGAGAGAGAGCCCTCCAGAAATTTGAAGCC 1680  
 DB 1646 CAATGTAATGGCCGCTTAAATGATGATGAGAGAGAGAGCCCTCCAGAAATTTGAAGCC 1705

QY 1681 CGCCCCCTTGGGCGGCGGAGACATGATGATCCCTCCAGTCAACATGGGCTTTATG 1740  
 DB 1706 CGCCCCCTTGGGCGGCGGAGACATGATGATCCCTCCAGTCAACATGGGCTTTATG 1765  
 QY 1741 GTCAAGATGTCATATGCTTGGCTGCGCTGCTACCGATTA 1779  
 DB 1766 GTCAAGATGTCATATGCTTGGCTGCGCTGCTACCGATTA 1804  
 RESULT 4  
 ID AAH22671  
 AAH22671 standard; DNA; 2636 BP.  
 XX  
 AC AAH22671;  
 XX  
 DT 07-SEP-2001 (first entry)  
 XX  
 DE Heparanase-like protein Hpa2 splice variant #1 encoding DNA.  
 XX  
 KM Heparanase; splice variant; homologue; heparanase-like protein; Hpa2;  
 KM cytosolic; neuroprotective; cerebroprotective; immunosuppressive;  
 KM antiproliferative; neurotropic; antiinflammatory; antiarthritic; antiaesthetic;  
 KM antidiabetic; antiarteriosclerotic; vulnerrary; ds.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 5'UTR 1..600  
 FT /\*tag= a  
 FT CDS 601..2379  
 FT /\*tag= b  
 FT /transl\_except= {pos: 1309..1311, aa: Xaa}  
 FT /note= "Xaa = unknown"  
 FT 3'UTR 2377..2636  
 FT /\*tag= c  
 PN WO200146392-A2.  
 PD 28-JUN-2001.  
 XX  
 PF 21-DEC-2000; 2000WO-GB004963.  
 XX  
 PR 22-DEC-1999; 99GB-0030392.  
 PR 07-APR-2000; 2000GB-00008713.  
 XX  
 PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
 PI McKenzie EA, Stamps AC, Terrett JA, Tyson KL;  
 XX  
 DR WPI; 2001-418056/44.  
 DR P-PSDB; AAB85215.  
 XX  
 PT Novel homologue of heparanase, present in three splice variants, useful  
 PT for identifying agents that modulate heparanase, useful in the treatment  
 PT and/or prophylaxis of abnormal levels of heparanase.  
 PS Claim 12; Fig 1; 97pp; English.  
 XX  
 CC The invention provides a homologue to heparanase which is present in  
 CC three splice variants. The heparanase homologue polypeptide is useful in  
 CC the treatment of a human or non-human animal or for use in diagnosis.  
 CC Vectors comprising the heparanase homologue polynucleotides are useful in  
 CC the transformation or transfection of a prokaryotic or eukaryotic host.  
 CC The modulators of the polypeptide are useful in the manufacture of a  
 CC medicament for the treatment and/or prophylaxis of a condition/disease  
 CC associated with abnormal levels of the heparanase homologue, including  
 CC cancer, central nervous system (CNS) and neurodegenerative diseases,  
 CC cardiovascular diseases such as restenosis following angioplasty and  
 CC atherosclerosis, autoimmune diseases, psoriasis, lupus erythematosus,  
 CC allografts, inflammatory diseases, arthritis, vascular restenosis, tumour  
 CC growth and progression, asthma, Alzheimer's disease, diabetic  
 CC retinopathy, wound healing and inflammation. The polypeptide is also  
 CC useful in diagnosis and research. The present sequence represents the

CC nucleotide sequence of the largest splice variant of the heparanase-like  
 CC protein Hpa2 of the invention  
 XX

Sequence 2636 BP; 707 A; 662 C; 673 G; 591 T; 0 U; 3 Other;

Query Match 99.7%; Score 1773.8; DB 4; Length 2636;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1775; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATAGAGGTGGTGTGTGCTTCCCTGGAAGCATGCTCCAGCAACTCCGCGCCCGCGG 60  
 Db ATAGAGGTGGTGTGTGCTTCCCTGGAAGCATGCTCCAGCAACTCCGCGCCCGCGG 660  
 QY 61 TGGCTTGAAGCCCGGAGGCTCTTCACTTGGCTTGTGCTCCATCTCTCCCTTCTCCAG 120  
 Db TGGCTTGAAGCCCGGAGGCTCTTCACTTGGCTTGTGCTCCATCTCTCCCTTCTCCAG 720  
 QY 121 GCTGAG 180  
 Db GCTGAG 780  
 QY 721 GCTGAG 780  
 Db GCTGAG 780  
 QY 181 ATTCTAATTGATGAG 240  
 Db ATTCTAATTGATGAG 840  
 QY 781 ATTCTAATTGATGAG 840  
 Db ATTCTAATTGATGAG 840  
 QY 241 CTGAGAGTGGATCCGCTCCATCATTCATGATGCTGCTGATTTCTTAAGCTCCAGAGCG 300  
 Db CTGAGAGTGGATCCGCTCCATCATTCATGATGCTGCTGATTTCTTAAGCTCCAGAGCG 900  
 QY 841 CTGAGAGTGGATCCGCTCCATCATTCATGATGCTGCTGATTTCTTAAGCTCCAGAGCG 900  
 Db CTGAGAGTGGATCCGCTCCATCATTCATGATGCTGCTGATTTCTTAAGCTCCAGAGCG 900  
 QY 301 TTGGTGAAGCCCTGGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
 Db TTGGTGAAGCCCTGGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960  
 QY 901 TTGGTGAAGCCCTGGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960  
 Db TTGGTGAAGCCCTGGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960  
 QY 361 GATCTTCTGAGAGTTCAG 420  
 Db GATCTTCTGAGAGTTCAG 420  
 QY 421 GATCTTCTGAGAGTTCAG 480  
 Db GATCTTCTGAGAGTTCAG 1020  
 QY 1021 GATCTTCTGAGAGTTCAG 1080  
 Db GATCTTCTGAGAGTTCAG 1080  
 QY 481 CAG 540  
 Db CAG 540  
 QY 1081 CAG 1140  
 Db CAG 1140  
 QY 541 GAGAGTCAAGTGAATGCTGCTTCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600  
 Db GAGAGTCAAGTGAATGCTGCTTCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200  
 QY 1141 GAGAGTCAAGTGAATGCTGCTTCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200  
 Db GAGAGTCAAGTGAATGCTGCTTCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200  
 QY 601 ATATTAACAG 660  
 Db ATATTAACAG 1260  
 QY 1201 ATATTAACAG 1260  
 Db ATATTAACAG 1260  
 QY 661 CTGATATTTGCTCTTAATGAG 720  
 Db CTGATATTTGCTCTTAATGAG 1320  
 QY 1261 CTGATATTTGCTCTTAATGAG 1320  
 Db CTGATATTTGCTCTTAATGAG 1320  
 QY 721 GCGCTGAAGTGTGTAAG 780  
 Db GCGCTGAAGTGTGTAAG 1380  
 QY 1321 GCGCTGAAGTGTGTAAG 1380  
 Db GCGCTGAAGTGTGTAAG 1380  
 QY 781 AATGAGCCAAATTAATGAG 840  
 Db AATGAGCCAAATTAATGAG 1440  
 QY 1381 AATGAGCCAAATTAATGAG 1440  
 Db AATGAGCCAAATTAATGAG 1440  
 QY 841 AAGGATTATCAATCAAGTGAAG 900  
 Db AAGGATTATCAATCAAGTGAAG 1500  
 QY 1441 AAGGATTATCAATCAAGTGAAG 1500  
 Db AAGGATTATCAATCAAGTGAAG 1500  
 QY 901 TTATATGAGCCCTTAATTTGGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960  
 Db TTATATGAGCCCTTAATTTGGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560

QY 961 ATGAGAGTGGCAGAGAGATCAATGAG 1020  
 Db ATGAGAGTGGCAGAGAGATCAATGAG 1620  
 QY 1561 ATGAGAGTGGCAGAGAGATCAATGAG 1620  
 Db ATGAGAGTGGCAGAGAGATCAATGAG 1620  
 QY 1021 CGGAGTGGTCAAGGATGAG 1080  
 Db CGGAGTGGTCAAGGATGAG 1680  
 QY 1621 CGGAGTGGTCAAGGATGAG 1680  
 Db CGGAGTGGTCAAGGATGAG 1680  
 QY 1081 ATTAGAGAAAATTCAGAAAGTGTAAATACATACATCACTCCAGAGAAAGAGAGAGAGAGAGAGAG 1140  
 Db ATTAGAGAAAATTCAGAAAGTGTAAATACATACATCACTCCAGAGAAAGAGAGAGAGAGAGAGAG 1740  
 QY 1681 ATTAGAGAAAATTCAGAAAGTGTAAATACATACATCACTCCAGAGAAAGAGAGAGAGAGAGAGAG 1740  
 Db ATTAGAGAAAATTCAGAAAGTGTAAATACATACATCACTCCAGAGAAAGAGAGAGAGAGAGAGAG 1740  
 QY 1141 GGTGAGTGAACCACTGAG 1200  
 Db GGTGAGTGAACCACTGAG 1800  
 QY 1741 GGTGAGTGAACCACTGAG 1800  
 Db GGTGAGTGAACCACTGAG 1800  
 QY 1201 TTCTTATGTTGAACAATTAAG 1260  
 Db TTCTTATGTTGAACAATTAAG 1860  
 QY 1801 TTCTTATGTTGAACAATTAAG 1860  
 Db TTCTTATGTTGAACAATTAAG 1860  
 QY 1261 CACTCATTTTTTGAACATGAGATACATCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320  
 Db CACTCATTTTTTGAACATGAGATACATCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920  
 QY 1861 CACTCATTTTTTGAACATGAGATACATCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920  
 Db CACTCATTTTTTGAACATGAGATACATCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920  
 QY 1321 GACTACTGAGCTCTCTCTCTCTCAAG 1380  
 Db GACTACTGAGCTCTCTCTCTCTCAAG 1980  
 QY 1921 GACTACTGAGCTCTCTCTCTCTCAAG 1980  
 Db GACTACTGAGCTCTCTCTCTCTCAAG 1980  
 QY 1381 GTGAGTGGGCTCCAG 1440  
 Db GTGAGTGGGCTCCAG 2040  
 QY 1981 GTGAGTGGGCTCCAG 2040  
 Db GTGAGTGGGCTCCAG 2040  
 QY 1441 TATGCTGAG 1500  
 Db TATGCTGAG 2100  
 QY 2041 TATGCTGAG 2100  
 Db TATGCTGAG 2100  
 QY 1501 ATCATCAACTTGCATGATCAAG 1560  
 Db ATCATCAACTTGCATGATCAAG 2160  
 QY 2101 ATCATCAACTTGCATGATCAAG 2160  
 Db ATCATCAACTTGCATGATCAAG 2160  
 QY 1561 CTGAGTTCACAGATGAG 1620  
 Db CTGAGTTCACAGATGAG 2220  
 QY 2161 CTGAGTTCACAGATGAG 2220  
 Db CTGAGTTCACAGATGAG 2220  
 QY 1621 CAAGTGAATGAG 1680  
 Db CAAGTGAATGAG 2280  
 QY 2221 CAAGTGAATGAG 2280  
 Db CAAGTGAATGAG 2280  
 QY 1681 GCGCCCTTGGGAGCGGAG 1740  
 Db GCGCCCTTGGGAGCGGAG 2340  
 QY 2281 GCGCCCTTGGGAGCGGAG 2340  
 Db GCGCCCTTGGGAGCGGAG 2340  
 QY 1741 GTCAAGATGCAATGCTTTGGAG 1779  
 Db GTCAAGATGCAATGCTTTGGAG 2379  
 QY 2341 GTCAAGATGCAATGCTTTGGAG 2379  
 Db GTCAAGATGCAATGCTTTGGAG 2379

# RESULT 5

AAD29202  
 ID AAD29202 standard; cDNA; 1847 BP.

XX AAD29202;

XX 07-MAY-2002 (first entry)

XX Human heparanase-2AB splice variant cDNA.

XX Human; heparanase-2AB; Hep-2; wound healing; angiogenesis; resection;

XX atherosclerosis; neurodegenerative disease; inflammation; proteinase;

XX viral infection; autoimmune lesion; renal failure; pancreatic cancer;

XX dystrophic muscular disease; heart disease; gene therapy; enzyme; ss.

XX Homo sapiens.



Key	Location/Qualifiers
FT	85..1833
FT	/*tag=a
FT	/product= "Human heparanase-2AB splice variant protein"
XX	
XX	MO300204645-A2.
XX	
XX	17-JAN-2002.
XX	
XX	12-JUL-2001; 2001MO-EP080894.
XX	
XX	12-JUL-2000; 2000EP-00202442.
XX	
XX	(VLA-A) VLAAMS INTERNUNIVERSITAIR INST BIOTECNOG.
XX	
XX	David G, Duerri J;
XX	
XX	WPI; 2002-171719/22.
XX	P-PDB; AAE18326.
XX	
XX	Heparanase-2 polypeptides and polynucleotides, useful for useful in wound healing, angiogenesis, and for treating restenosis, atherosclerosis, inflammation, neurodegenerative diseases, and viral infections.
XX	
XX	Claim 2; Page 35-38; 54pp; English.
XX	
XX	The invention relates to human heparanase-2 (Hep-2) polypeptides and polynucleotides. Heparanase-2 protein is useful in wound healing, angiogenesis, restenosis, atherosclerosis, neurodegenerative diseases, inflammation and viral infections, as well as in neutralising plasma heparin as a potential replacement of protamine. Antiheparanase-2 antibiotics may be used for immunodetection and diagnosis of micrometastases, autoimmune lesions, renal failure in biopsy specimens, plasma samples and body fluids. Molecules, which can agonise or antagonise heparanase 2 catalytic activity may also be used as a medicament. Polymorphisms in the polynucleotide sequence are useful in the identification of individuals having a predisposition to acquire diseases resulting from a increased or decreased expression of their activity. Such molecules can be used to treat pancreatic cancer, dys trophic muscular diseases and or heart diseases. Polynucleotides of the invention are used in gene therapy. The present sequence is human heparanase-2AB splice variant CDNA
XX	
XX	Sequence 1847 BP; 465 A; 491 C; 456 G; 434 T; 0 U; 1 Other;
XX	
XX	Query Match 99.4%; Score 1769; DB 6; Length 1847;
XX	Best Local Similarity 99.6%; Pred. No. 0;
XX	Matches 1772; Conservative 1; Mismatches 6; Indels 0; Gaps 0
XX	
XX	1 ATGAGAGGAGCTTTGGAGCTTCCCTTAAGGACGAGCCCTCAGAGAACTCCGCCCGCCGCG 60
XX	55 ATGAGGAGGCTTTGGAGCTTCCCTTAAGGACGAGCCCTCAGAGAACTCCGCCCGCCGCG 114
XX	
XX	61 TGCCTAGCCCCCGGGGGCTCTCACTTGGCTCTGTGCTTCATCTCTCCTTTCTCCAG 120
XX	115 TGCCTAGCCCCCGGGGGCTCTCACTTGGCTCTGTGCTTCATCTCTCCTTTCTCCAG 174
XX	
XX	121 GCTGAGACAGAGAGACCTTGGCTCTTACACAGAGCTGACGTTTGAAGGAAAGACCTTG 180
XX	175 GCTGAGACAGAGAGACCTTGGCTCTTACACAGAGCTGACGTTTGAAGGAAAGACCTTG 234
XX	
XX	181 ATTCTACTTGAATGTAGACACCAAGAACCAATGACAGCTCAATGAGAACTTCTCTCT 240
XX	235 ATTCTACTTGAATGTAGACACCAAGAACCAATGACAGCTCAATGAGAACTTCTCTCT 294
XX	
XX	241 CTGACAGCTGATCCCTTCATCATTTATGATGCTGGCTTCATTTCTTAAGCTCAAGCGC 300
XX	295 CTGACAGCTGATCCCTTCATCATTTATGATGCTGGCTTCATTTCTTAAGCTCAAGCGC 354
XX	
XX	301 TTGGGAGACCCCTGGAGCCCGGGGAACTTTCGCCCCGCTTTCGCGCTTGGGGGCAAAAGAAC 360
XX	355 TTGGGAGACCCCTGGAGCCCGGGGAACTTTCGCCCCGCTTTCGCGCTTGGGGGCAAAAGAAC 414

[illegible]





```

QY 721 GCCCTGAGTGTGTTGAAGTACAGCCGACAAAAGTACACATTTCTTGGGAACCTGGGT 780
DB 584 GCCCTGAGTGTGTTGAAGTACAGCCGACAAAAGTACACATTTCTTGGGAACCTGGGT 643
QY 781 AATGAGCAATATATATATGAGCATGAGCATGAGCGGAGAGTAAATGGAGCCAGTTGGGA 840
DB 644 AATGAGCAATATATATGAGCATGAGCATGAGCGGAGAGTAAATGGAGCCAGTTGGGA 703
QY 841 AAGGATTATACCACTGAGTAAAGCCTTGTGACGCCCAATCCGATTATTTCCAGAGCCAGC 900
DB 704 AAGGATTATACCACTGAGTAAAGCCTTGTGACGCCCAATCCGATTATTTCCAGAGCCAGC 763
QY 901 TTATATGAGCCCTATATATTTGGGCGGCGAGAGAGATGTATGAGCCCTCCATGATGATTC 960
DB 764 TTATATGAGCCCTATATTTGGGCGGCGAGAGAGATGTATGAGCCCTCCATGATGATTC 823
QY 961 ATGAAAGTGGCAGAGAGTACAGTATGATGACATTACCTGGCAATGCTTACATTGATGAGC 1020
DB 824 ATGAAAGTGGCAGAGAGTACAGTATGATGACATTACCTGGCAATGCTTACATTGATGAGC 883
QY 1021 CGGCTGGTCAAGGTGATGAGCCTTCCGAAAACCTGCTGTGATGACACACTCTGACCCAG 1080
DB 884 CGGCTGGTCAAGGTGATGAGCCTTCCGAAAACCTGCTGTGATGACACACTCTGACCCAG 943
QY 1081 ATTAGAAAAATTCAGAAAGTGTATATATACATCTCCAGAAAAGAAATTTGGCTTGA 1140
DB 944 ATTAGAAAAATTCAGAAAGTGTATATATACATCTCCAGAAAAGAAATTTGGCTTGA 1003
QY 1141 GGTGTGTGACCACTGATGAGGAGCAAAATCTATCCGATTCCTATGTTGTGAGGA 1200
DB 1004 GGTGTGTGACCACTGATGAGGAGCAAAATCTATCCGATTCCTATGTTGTGAGGA 1063
QY 1201 TTCTTATGTTGAACTTTAGATGCTGGCAATCAGGGCATTTGATGATACCG 1260
DB 1064 TTCTTATGTTGAACTTTAGATGCTGGCAATCAGGGCATTTGATGATACCG 1123
QY 1261 CACTCATTTTTTGAACCATGATACATCACTGTGTGACCAAAATTTTAAACCATTAACA 1320
DB 1124 CACTCATTTTTTGAACCATGATACATCACTGTGTGACCAAAATTTTAAACCATTAACA 1183
QY 1321 GACTACTGGCTCTCTCTCTCTCTCAAGCGCTGTGAGCGCCCAAGCTTGGCTGTGAT 1380
DB 1184 GACTACTGGCTCTCTCTCTCTCTCTCAAGCGCTGTGAGCGCCCAAGCTTGGCTGTGAT 1243
QY 1381 GTGGCTGGGCTCCAGCGGAAGCCACGGCTGGCCGAGTGTCCGGGACAAATTAAGAT 1440
DB 1244 GTGGCTGGGCTCCAGCGGAAGCCACGGCTGGCCGAGTGTCCGGGACAAATTAAGAT 1303
QY 1441 TATGCTACTGCAAAACCAACAACACACACATGATGTTGGGTCATTAACATTTT 1500
DB 1304 TATGCTACTGCAAAACCAACAACACACACATGATGTTGGGTCATTAACATTTT 1363
QY 1501 ATCATCAACTGATCATGATCAAGAAAGAAATCAAGTGTGCTGGGACTCTCAGAGCAAG 1560
DB 1364 ATCATCAACTGATCATGATCAAGAAAGAAATCAAGTGTGCTGGGACTCTCAGAGCAAG 1423
QY 1561 CTGGTTACACAGTACCTGTGAGCGCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620
DB 1424 CTGGTTACACAGTACCTGTGAGCGCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1483
QY 1621 CAATGAAATGAGCCCTTATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
DB 1484 CAATGAAATGAGCCCTTATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1543
QY 1681 CGCCCCCTTGGGCGGCGGAGCAATGATGATGATGATGATGATGATGATGATGATG 1740
DB 1544 CGCCCCCTTGGGCGGCGGAGCAATGATGATGATGATGATGATGATGATGATGATG 1603
QY 1741 GTCAAGATGTCAATGCTTTGGCTGTGCGCTACCGATTA 1779
DB 1604 GTCAAGATGTCAATGCTTTGGCTGTGCGCTACCGATTA 1642

```

```

RESULT 8
AAB29204
ID AAB29204 standard; cDNA; 1685 BP.
AC AAB29204;
XX
XX 07-MAY-2002 (first entry)
DT
XX
XX Human heparanase-2A splice variant cDNA.
DE
XX
XX Human; heparanase-2A; Hep-2; wound healing; angiogenesis; restenosis;
KW atherosclerosis; neurodegenerative disease; inflammation; prolamine;
KW viral infection; autoimmune lesion; renal failure; pancreatic cancer;
KW dystrophic muscular disease; heart disease; gene therapy; enzyme; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 85.1671
FT CDS /*tag= a
FT product= "Human heparanase-2A splice variant protein"
FT
XX
XX MO200204645-A2.
XX
XX 17-JAN-2002.
XX
XX 12-JUL-2001; 2001WO-EP008094.
XX
XX 12-JUL-2000; 2000EP-00202442.
XX
XX (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
XX David G. Duerr J;
XX
XX WPI: 2002-171719/22.
XX
XX P-PSDB; AAE18327.
XX
XX
XX Heparanase-2 polypeptides and polynucleotides, useful for useful in wound
PT healing, angiogenesis, and for treating restenosis, atherosclerosis,
PT inflammation, neurodegenerative diseases, and viral infections.
XX
XX
XX Disclosure; Page 42-45; 54pp; English.
XX
XX The invention relates to human heparanase-2 (Hep-2) polypeptides and
CC polynucleotides. Heparanase-2 protein is useful in wound healing,
CC angiogenesis, restenosis, atherosclerosis, neurodegenerative diseases,
CC inflammation and viral infections, as well as in neutralising plasma
CC heparin as a potential replacement of protamine. Antiheparanase-2
CC antibodies may be used for immunodetection and diagnosis of
CC micrometastases, autoimmune lesions, renal failure in biopsy specimens,
CC plasma samples and body fluids. Molecules, which can agonise or
CC antagonise heparanase 2 catalytic activity may also be used as a
CC medicament. Polymorphisms in the polynucleotide sequence are useful in
CC the identification of individuals having a predisposition to acquire
CC diseases resulting from an increased or decreased expression of their
CC activity. Such molecules can be used to treat pancreatic cancer,
CC dystrophic muscular diseases and/or heart diseases. Polynucleotides of
CC the invention are used in gene therapy. The present sequence is human
CC heparanase-2A splice variant cDNA
XX
XX
SQ Sequence 1685 BP; 415 A; 456 C; 419 G; 394 T; 0 U; 1 Other;
Query Match 81.0%; Score 1441.4; DB 6; Length 1685;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 1614; Conservative 1; Mismatches 2; Indels 162; Gaps 1;
QY 1 ATGAGGAGTGTGCTTGTGCTTCCCTGGAAGCCATGCGCTCCAGCACTCCGCGCCCGCGG 60
DB 55 ATGAGGAGTGTGCTTGTGCTTCCCTGGAAGCCATGCGCTCCAGCACTCCGCGCCCGCGG 114
QY 61 TGCCTAGCCCGGAGGAGCTCTACTTGGCTGTGTGCTGCATCTCTCCCTTCTCCAG 120
DB 115 TGCCTAGCCCGGAGGAGCTCTACTTGGCTGTGTGCTGCATCTCTCCCTTCTCCAG 174

```

```

Qy 121 GCTGAGACAGAGACCCCTTGGCTGTAGACAGAGCTGCAGGTTTGAAGAAAAGACCTG 180
Db 175 GCTGAGACAGAGACCCCTTGGCTGTAGACAGAGCTGCAGGTTTGAAGAAAAGACCTG 234
Qy 181 ATTCTACTGATGATGAGACACAGAAACCCAGTCAAGACAGTCAATGAGAACTTCTCTCT 240
Db 235 ATTCTACTGATGATGAGACACAGAAACCCAGTCAAGACAGTCAATGAGAACTTCTCTCT 294
Qy 241 CTGACAGCTGATCCGCTCATTCATTCATGATGATGATGATGATGATGATGATGATGATG 300
Db 295 CTGACAGCTGATCCGCTCATTCATTCATGATGATGATGATGATGATGATGATGATGATG 354
Qy 301 TTGGTGAACCTGGCCCGGGGACCTTGGCCCGCTTCTGCGCTGGGGGGGCAAAAGAAC 360
Db 355 TTGGTGAACCTGGCCCGGGGACCTTGGCCCGCTTCTGCGCTGGGGGGGCAAAAGAGGC 414
Qy 361 GACTTCCTGAGTTCAGAACCTGAGAACCCGGCGAAAGCCGGCGGGCCCGGGCCCG 420
Db 415 GACTTCCTGAGTTCAGAACCTGAGAACCCGGCGAAAGCCGGCGGGCCCGGGCCCG 474
Qy 421 GATTACTATCTCAAAACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db 475 GATTACTATCTCAAAACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 501
Qy 481 CAGAAAGGCTGCAAGATGCGCCAGACCCCTGATGATGATGATGATGATGATGATGATG 540
Db 502 ----- 501
Qy 541 GCAGCTCAAGTGCATCGTCTTCTTAAAGAGCAATCTCCAAATCTTAAGATATCTC 600
Db 502 ----- 501
Qy 601 ATATTAACAGCCAGTCTCTAGACAACTTATTAATTGCTGATGCTGAGACTCCAC 660
Db 502 -----GCCAGCTCTAGACAACTTATTAATTGCTGATGCTGAGACTCCAC 552
Qy 661 CTGATATTTGCTCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db 553 CTGATATTTGCTCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 612
Qy 721 GCCCTAGCTGTTGAAGTACAGCGCCAGCAAAATCTTAAGCAATCTTGGGAGCTGGGT 780
Db 613 GCCCTAGCTGTTGAAGTACAGCGCCAGCAAAATCTTAAGCAATCTTGGGAGCTGGGT 672
Qy 781 AATGAGCAAAATTAATCTATCGACCATGATGCGCGGAGATTAATGAGCAAGTTGGGA 840
Db 673 AATGAGCAAAATTAATCTATCGACCATGATGCGCGGAGATTAATGAGCAAGTTGGGA 732
Qy 841 AAGGATTAATCTATCGACCATGATGCGCGGAGATTAATGAGCAAGTTGGGA 900
Db 733 AAGGATTAATCTATCGACCATGATGCGCGGAGATTAATGAGCAAGTTGGGA 792
Qy 901 TTATATGACCTTAATATTTGGGCGCGGAGAAATGATGATGATGATGATGATGATGATG 960
Db 793 TTATATGACCTTAATATTTGGGCGCGGAGAAATGATGATGATGATGATGATGATGATG 852
Qy 961 ATGAGGTGACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
Db 853 ATGAGGTGACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 912
Qy 1021 CGGGTGGTCAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
Db 913 CGGGTGGTCAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 972
Qy 1081 ATTAGGAAAATTCAGAAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
Db 973 ATTAGGAAAATTCAGAAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1032
Qy 1141 GGTGTGTGACACCTGAGCTGAGGAGCAAAATATCTATCCATTCCTATGCTGAGAGA 1200
Db 1033 GGTGTGTGACACCTGAGCTGAGGAGCAAAATATCTATCCATTCCTATGCTGAGAGA 1092

```

```

Qy 1201 TTCTTATGTTGAACACTTAGAATGCTGCGCAATCAGGGCATGATGATGATGATGATG 1260
Db 1093 TTCTTATGTTGAACACTTAGAATGCTGCGCAATCAGGGCATGATGATGATGATGATG 1152
Qy 1261 CACTGATTTTTTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
Db 1153 CACTGATTTTTTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1212
Qy 1321 GACTACTGCGCTCTCTCTCTCTCAAGCGCTGATGCGGCCCAAGCTTGGCTGATGAT 1380
Db 1213 GACTACTGCGCTCTCTCTCTCTCAAGCGCTGATGCGGCCCAAGCTTGGCTGATGAT 1272
Qy 1381 GTGGCTGGGCTCCAGCGGAGCCAGCGGCTGCGGATGATGCGGAGCAAACTPAAGATT 1440
Db 1273 GTGGCTGGGCTCCAGCGGAGCCAGCGGCTGCGGATGATGCGGAGCAAACTPAAGATT 1332
Qy 1441 TATGCTCACTGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1500
Db 1333 TATGCTCACTGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1392
Qy 1501 ATCATCAACTTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
Db 1393 ATCATCAACTTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1452
Qy 1561 CTGGTTACCAAGTACCTGCTGACGCTTATGCGGAGAGAGGCTTAAAGTCAAGTCA 1620
Db 1453 CTGGTTACCAAGTACCTGCTGACGCTTATGCGGAGAGAGGCTTAAAGTCAAGTCA 1512
Qy 1621 CACTGAATGCGGCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Db 1513 CACTGAATGCGGCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1572
Qy 1681 CGCCCGCTTGGGCGGCGGAGCAATGATGATGATGATGATGATGATGATGATGATGAT 1740
Db 1573 CGCCCGCTTGGGCGGCGGAGCAATGATGATGATGATGATGATGATGATGATGATGAT 1632
Qy 1741 GTCAAGAATGCAATGCTTGGCGCTGCGGATGATGATGATGATGATGATGATGATG 1779
Db 1633 GTCAAGAATGCAATGCTTGGCGCTGCGGATGATGATGATGATGATGATGATGATG 1671

RESULT 9
AB211527
ID AB211527 standard; cDNA; 1891 BP.
XX
AC AB211527;
XX
DT 20-JAN-2003 (first entry)
XX
DE Human polynucleotide seq ID NO 409.
XX
KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
KW cell-proliferative disorder; neurodegenerative disease; bacterial;
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
KW arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
KW antiarthritic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200270539-A2.
XX
PD 12-SEP-2002.
XX
PF 05-MAR-2002; 2002MO-US005095.
XX
PR 05-MAR-2001; 2001US-00799451.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;

```

PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,  
 PI Wehman T, Wang J, Wang D, Drmanac RT;  
 DR WPI; 2002-759812/82.  
 XX P-PSDB; ABP69310.

PT New polynucleotides comprising sequences assembled from expressed  
 PT sequence tags (ESTs), useful for treating cell-proliferative,  
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet  
 PT or coagulation disorders.

XX Claim 1; SEQ ID NO 409; 1012bp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide (1) comprising a  
 CC nucleotide sequence selected from any of 948 sequences (AB21113-  
 CC AB212066) or their mature protein coding portion, active domain coding  
 CC protein or complementary sequences. The polynucleotides are useful for  
 CC identifying expressed genes or for physical mapping of human genome. The  
 CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight  
 CC markers, as a food supplement, for generating antibodies, in medical  
 CC imaging, screening and diagnostic assays and for treating cell-  
 CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's  
 CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,  
 CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,  
 CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver  
 CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),  
 CC arthritis, etc. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/publicshed\_pct\_sequences  
 XX

SQ Sequence 1891 BP; 483 A; 513 C; 456 G; 439 T; 0 U; 0 Other;

Query Match 79.7%; Score 1417.8; DB 6; Length 1891;  
 Best Local Similarity 90.1%; Pred. No. 0;  
 Matches 1603; Conservative 0; Mismatches 2; Indels 174; Gaps 1;

QY 1 ATGAGGGTGGTCTGCTGCTTCCCTGAGGAGCATGCTCCAGCAACTCCCGCCCGCG 60  
 DB 20 ATGAGGGTGGTCTGCTGCTTCCCTGAGGAGCATGCTCCAGCAACTCCCGCCCGCG 79  
 QY 61 TGGCTTACCCCGGGGGCTCTTACTTGGCTCTGCTTCCATCTCTCTTCTCCAG 120  
 DB 80 TGGCTTACCCCGGGGGCTCTTACTTGGCTCTGCTTCCATCTCTCTTCTCCAG 139  
 QY 121 GCTGAGACGAGAGACCTTGGCTCTGAGACAGAGCTGAGGTTGAAGAAAGACCTG 180  
 DB 140 GCTGAGACGAGAGACCTTGGCTCTGAGACAGAGCTGAGGTTGAAGAAAGACCTG 199  
 QY 181 ATTCTATTGATGAGACCAAGAACCCAGTCAGACAGTCATGAGAACTTCTCT 240  
 DB 200 ATTCTATTGATGAGACCAAGAACCCAGTCAGACAGTCATGAGAACTTCTCT 259  
 QY 241 CTGAGCTGATCCCTGATCATTCATGATGAGTGGCTGATTTCTTAAGCTCCAGAGCG 300  
 DB 260 CTGAGCTGATCCCTGATCATTCATGATGAGTGGCTGATTTCTTAAGCTCCAGAGCG 319  
 QY 301 TTGGTGAACCTGAGCCCGGGAGCTTGGCCCGCTTCTGCGCTTCCGCGGCAAAAGAGAC 360  
 DB 320 TTGGTGAACCTGAGCCCGGGAGCTTGGCCCGCTTCTGCGCTTCCGCGGCAAAAGAGAC 379  
 QY 361 GACTTCTGAGATTCAGAACTTGAAGAACCCGCGGAGAGCCGCGGCGCG 420  
 DB 380 GACTTCTGAGATTCAGAACTTGAAGAACCCGCGGAGAGCCGCGGCGCG 439  
 QY 421 GATTACTATCTCAAAAATCTAGAGATGATGATTTGAGATGATTTGCTTGAATAA 480  
 DB 440 GATTACTATCTCAAAAATCTAGAGATGATGATTTGAGATGATTTGCTTGAATAA 499  
 QY 481 CAGAAAGGCTGCAAGATGCGCCGACACCTGATGATTTGCTGCTCCAAAGAGAG 540  
 DB 500 CAGAAAGGCTGCAAGATGCGCCGACACCTGATGATTTGCTGAGGCTCCAAAGAGAG 559  
 QY 541 GCAGCTCAGATGATGCTGTTCTTAAAGAGCAATTCCTCAATACTTCACTATCTC 600

DB 560 GCAGCTCAGATGATGCTGTTCTTAAAGAGCAATTCCTCAATACTTCACTATCTC 619  
 QY 601 ATATTAACAGCCAGGCTCTGAGCAAACTTTAACTTTGCTGATGCTGAGACTCCAC 660  
 DB 620 ATATTAACA----- 628  
 QY 661 CTGATATTTGCTCTAAATGCACTGCGTGAATCCCAATTACTCTGGAACAGTTCTAGT 720  
 DB 629 ----- 628  
 QY 721 GCCCTGAGTCTGTTGAATGACAGCCGCAAGAAAAGTAAACAATTCTTGGGAACCTGGGT 780  
 DB 629 ----- 628  
 QY 781 AATGAGCCAAATATATTCGAGACCATGATGCGCGGAGATAATGCAAGCCAGTTGGAA 840  
 DB 629 ---GAGCCAAATATATTCGAGACCATGATGCGCGGAGATAATGCAAGCCAGTTGGAA 685  
 QY 841 AAGGATTAATCCAGCTGAAAGAGCTGTTGACCCCATCCGATTTATTCAGAGCCAGC 900  
 DB 686 AAGGATTAATCCAGCTGAAAGAGCTGTTGACCCCATCCGATTTATTCAGAGCCAGC 745  
 QY 901 TTATATGAGCCCTATATTTGGAGCGGCGAGAGAAATGTCATGCGCCCTCTAGATGATTC 960  
 DB 746 TTATATGAGCCCTATATTTGGAGCGGCGAGAGAAATGTCATGCGCCCTCTAGATGATTC 805  
 QY 961 ATGAGAGTGGCAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
 DB 806 ATGAGAGTGGCAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 865  
 QY 1021 CGGCTGATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080  
 DB 866 CGGCTGATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 925  
 QY 1081 ATTAGAATAATTCAGAAAGTGGTTAATATATATATATATATATATATATATATATATAT 1140  
 DB 926 ATTAGAATAATTCAGAAAGTGGTTAATATATATATATATATATATATATATATATATAT 985  
 QY 1141 GGTGTGTGACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200  
 DB 986 GGTGTGTGACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1045  
 QY 1201 TTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260  
 DB 1046 TTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1105  
 QY 1261 CACTCATTTTTTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320  
 DB 1106 CACTCATTTTTTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1165  
 QY 1321 GACTACTGCTCTCTCTCTCTCTCAAGAGGCTGATGCGCCCAAGTCTTGGCTGTCAT 1380  
 DB 1166 GACTACTGCTCTCTCTCTCTCTCAAGAGGCTGATGCGCCCAAGTCTTGGCTGTCAT 1225  
 QY 1381 GTGGCTGGCTCTCAAGCGAAAGCCAGGCTGCGGAGATTCGAGGCAAACTTAAGAT 1440  
 DB 1226 GTGGCTGGCTCTCAAGCGAAAGCCAGGCTGCGGAGATTCGAGGCAAACTTAAGAT 1285  
 QY 1441 TATGCTCACTGCAAAACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1500  
 DB 1286 TATGCTCACTGCAAAACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1345  
 QY 1501 ATCATCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560  
 DB 1346 ATCATCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1405  
 QY 1561 CTGCTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620  
 DB 1406 CTGCTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1465  
 QY 1621 CAAGTGAATGAGCCAGCCCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680



Db 1466 CAAGTGAATGGCCAGCCCTAGTATGATGAGACAGACGAGCCCTCCAGATTTAGAGCC. 1525  
 Qy 1661 GCGCCCTTTGGGCGCGCGGAGATTTGATCCCTCCAGTACACATGGGCTTTATGTG 1740  
 Db 1526 GCGCCCTTTGGGCGCGCGGAGATTTGATCCCTCCAGTACACATGGGCTTTATGTG 1585  
 Qy 1741 GTCAAGATGTCATGCTTTGGGCTGCGGTACCGATTA 1779  
 Db 1586 GTCAAGATGTCATGCTTTGGGCTGCGGTACCGATTA 1624

RESULT 10  
 AAI70705  
 ID AAI70705 standard; cDNA; 2326 BP.  
 XX  
 AC AAI70705;  
 XX  
 DT 04-FEB-2002 (first entry)  
 XX  
 XX Human prepro-heparanase II cDNA.  
 DE Heparanase II; human; vulnery; angiogenesis inhibitor;  
 XX antiinflammatory; cyrostatic; gene therapy; diagnosis; ss.  
 KM  
 XX Homo sapiens.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 XX CDS 25..1629  
 XX FT /tag= a  
 XX FT 25..147  
 XX FT sig\_peptide  
 XX FT 148..1626  
 XX FT /tag= b  
 XX FT mat\_peptide  
 XX FT 148..1626  
 XX FT /tag= c  
 XX  
 XX MO200181569-A2.  
 XX  
 XX 01-NOV-2001.  
 XX  
 XX 17-APR-2001; 2001MO-US010804.  
 XX  
 XX 20-APR-2000; 2000US-0199072P.  
 XX  
 XX (PHAA ) PHARMACIA & UPJOHN CO.  
 XX  
 XX Heinrichson RL, Bienkowski MJ;  
 XX  
 XX WPI; 2002-041402/05.  
 XX P-PSDB; AAM50337.  
 XX  
 XX Novel heparanase II polypeptide useful for identifying agents with alter  
 PT heparanase activity and for accelerating wound healing, blocking  
 PT angiogenesis or inflammation.  
 PS  
 PS Claim 2; Page 9-11; 65p; English.  
 XX  
 CC The present sequence is that of cDNA encoding novel human prepro-  
 CC heparanase II (see AAM50337), a paralogue of human heparanase I.  
 CC Identification of the cDNA involved a combination of protein sequencing  
 CC and mining of expressed sequence tag databases using human heparanase I  
 CC as query. Incyte clones 1634352 (prostate tumour), 3207353 (corpus  
 CC cavernosum) and 3704980 (corpus cavernosum) were detected, and additional  
 CC sequences were obtained from Incyte clones 3529440 (bladder), and 3385825  
 CC (oesophagus) and from 5' RACE products using human prostate, small  
 CC intestine, bladder or heart cDNA templates. Heparanase II is a secreted  
 CC protein that shows 43% identity at the amino acid level to heparanase I.  
 CC The prepro-protein encoded by the current sequence is processed to remove  
 CC a 41-amino acid leader peptide, and further processed to remove internal  
 CC amino acids, yielding the 8 kDa and 50 kDa subunits of the heparanase II  
 CC enzyme. Heparanase I and II have a non-overlapping expression pattern in  
 CC human tissues and each may serve tissue-specific functional roles. The  
 CC invention provides heparanase II nucleic acids, vectors, host cells,  
 CC polypeptides and antibodies. Heparanase II is useful for identifying an  
 CC agent that alters heparanase activity, and which can be used to treat a

CC disease state. Nucleic acids are useful for large scale expression of  
 CC heparanase II, in the design of probes and primers, and in the  
 CC identification of a genetic alteration(s) that underlies a disease state or  
 CC states. Inhibiting heparanase II activity is useful for treating or  
 CC preventing metastasis, cancer, CNS and neurodegenerative diseases,  
 CC inflammation and cardiovascular disease such as restenosis following  
 CC angioplasty and atherosclerosis. Heparanase II is useful for accelerating  
 CC wound healing, blocking angiogenesis, degradation of heparin and  
 CC neutralization of heparin's anticoagulant properties during surgery.  
 CC Heparanase or an agent that enhances heparanase activity can also be  
 CC infused into the vasculature to block accumulation and diapedesis of  
 CC neutrophils at sites of inflammation with or without added domains to  
 CC confer selectivity in delivery

XX  
 XX Sequence 2326 BP; 593 A; 631 C; 550 G; 552 T; 0 U; 0 Other;  
 SQ  
 Query Match 79.7%; Score 1417.8; DB 6; Length 2326;  
 Best Local Similarity 90.1%; Pred. No. 0;  
 Matches 1603; Conservative 0; Mismatches 2; Indels 174; Gaps 1;

Qy 1 ATGAGGCTGCTTTGCTTCTCTGCTTCTCTGAGCATGCTCCAGCAACCCGCCCCCGCG 60  
 Db 25 ATGAGGCTGCTTTGCTTCTCTGCTTCTCTGAGCATGCTCCAGCAACCCGCCCCCGCG 84  
 Qy 61 TGCCCTAGCCCCGGGGGCTCTCTACTTGGCTCTGTGCTCATCTCCCTTTCTCCAG 120  
 Db 85 TGCCCTAGCCCCGGGGGCTCTCTACTTGGCTCTGTGCTCATCTCCCTTTCTCCAG 144  
 Qy 121 GCTGAGACAGAGAGACCTTGTGCTGAGACAGAGCTGAGGCTTTGAAGAAAGACCTG 180  
 Db 145 GCTGAGACAGAGAGACCTTGTGCTGAGACAGAGCTGAGGCTTTGAAGAAAGACCTG 204  
 Qy 181 ATTCTACTTATGATGACACCAAGAACCCAGTACAGACGTCAATGAGAACTTCTCT 240  
 Db 205 ATTCTACTTATGATGACACCAAGAACCCAGTACAGACGTCAATGAGAACTTCTCT 264  
 Qy 241 CTGCAAGCTGATCCGTCATCATTCATGATGAGCTGAGCTGATTTCTTAAGCTCCAAAGCG 300  
 Db 265 CTGCAAGCTGATCCGTCATCATTCATGATGAGCTGAGCTGATTTCTTAAGCTCCAAAGCG 324  
 Qy 301 TTGGTGAACCTGAGCGCGGAGACTTGGCCGCTTTGCGCTTCCGAGGAGAAAGAGAC 360  
 Db 325 TTGGTGAACCTGAGCGCGGAGACTTGGCCGCTTTGCGCTTCCGAGGAGAAAGAGAC 384  
 Qy 361 GACTTCCGCACTTCCAGAACTGAGAGACCCGCGGAGAAAGCGCGCGGCGCG 420  
 Db 385 GACTTCCGCACTTCCAGAACTGAGAGACCCGCGGAGAAAGCGCGCGGCGCG 444  
 Qy 421 GATTACTATCTCAAAAATATGAGATGACATTTGGAAGTATTTGCTTATGATTA 480  
 Db 445 GATTACTATCTCAAAAATATGAGATGACATTTGGAAGTATTTGCTTATGATTA 504  
 Qy 481 CAGAAAGGCTGCAAGATGAGCCAGCAGCTGATGTATGCTGAGCTCCAAAGGAGAG 540  
 Db 505 CAGAAAGGCTGCAAGATGAGCCAGCAGCTGATGTATGCTGAGCTCCAAAGGAGAG 564  
 Qy 541 GCAAGTCAGATGATCTGTTCTTTTAAGAGCAATTTCCATTAATTAATTAATCTC 600  
 Db 565 GCAAGTCAGATGATCTGTTCTTTTAAGAGCAATTTCCATTAATTAATTAATCTC 624  
 Qy 601 ATATTAAGCAAGCTCTTGAAGCAAACTTATTAATTTGATGCTGAGACTCAC 660  
 Db 625 ATATTAAGCAAGCTCTTGAAGCAAACTTATTAATTTGATGCTGAGACTCAC 684  
 Qy 661 CTGATATTTGCTTAATGACATGCGGTGATCCCAATTAATCTCGAAGACATTTAGT 720  
 Db 684 CTGATATTTGCTTAATGACATGCGGTGATCCCAATTAATCTCGAAGACATTTAGT 744  
 Qy 721 GCGCTGAGCTCTGTTGAATGACAGCGCCAGCAAAAAGTACACATTTCTTGGAACTGGGT 780  
 Db 744 GCGCTGAGCTCTGTTGAATGACAGCGCCAGCAAAAAGTACACATTTCTTGGAACTGGGT 804  
 Qy 781 ATGAGCAAAATTAATTAATGACATGAGATGAGCGGAGCAATTAATGAGCAAGTTGGGA 840  
 Db 804 ATGAGCAAAATTAATTAATGACATGAGATGAGCGGAGCAATTAATGAGCAAGTTGGGA 868





```

QY 121 GCTGAGACAGAGAGACCTTGGCTGTGACAGAGCTGACAGTTTGAAGGAAAAGACCTTG 180
DB 721 GCTGAGACAGAGAGACCTTGGCTGTGACAGAGCTGACAGTTTGAAGGAAAAGACCTTG 780
QY 181 ATTCTACTGATGTGAGACCAAGAAACCAAGTCAGAGACAGTCATAGAGAACTTCTCTCT 240
DB 781 ATTCTACTGATGTGAGACCAAGAAACCAAGTCAGAGACAGTCATAGAGAACTTCTCTCT 840
QY 241 CTGACACTGGATCCGTCATCATTCATGATGAGTGGCTGATTTCTGTAGCTCAAGGCG 300
DB 841 CTGACACTGGATCCGTCATCATTCATGATGAGTGGCTGATTTCTGTAGCTCAAGGCG 900
QY 301 TTGGTACCTTGGCCCGGGGACCTTTCGCCCTTTCGCTGCGCTTGGGGGCAAAAGAAC 360
DB 901 TTGGTACCTTGGCCCGGGGACCTTTCGCCCTTTCGCTGCGCTTGGGGGCAAAAGAAC 960
QY 361 GACTTCTGCAAGTTCAGAAACCTGAGAGAACCCGGGCAAAAGCCGCGGGGGCCCGCG 420
DB 961 GACTTCTGCAAGTTCAGAAACCTGAGAGAACCCGGGCAAAAGCCGCGGGGGCCCGCG 1020
QY 421 GATTACTATCTCAAAAACCTATGAGATGACATTTGTCAGAGTGAATGTTGCTTAGATAA 480
DB 1021 GATTACTATCTCAAAAACCTATGAGATGACATTTGTCAGAGTGAATGTTGCTTAGATAA 1080
QY 481 CAGAAAGGCTGCAAGATTGCCGACACCTGATGTTATGCTGCTGCCAAGGGAGAG 540
DB 1081 CAGAAAGGCTGCAAGATTGCCGACACCTGATGTTATGCTGAGCTCCAAAGGGAGAG 1140
QY 541 GCAGCTCAAGATGATCGGTTCTTCTAAAGAGACATTTCCCAATACAGTAATCTC 600
DB 1141 GCAGCTCAAGATGATCGGTTCTTCTAAAGAGACATTTCCCAATACAGTAATCTC 1200
QY 601 AATATTAACGACGAGTCTCTAGACAAACCTTATATTAATCTTGTGATGCTCTGACCTCAC 660
DB 1201 AATATTAAC----- 1209
QY 661 CTGATATTTTGTCTTAATGACATGCGGTCCGTAATCCCAATACCTCTGAGAACATTTCTAGT 720
DB 1210 ----- 1209
QY 721 GCCCTGAGTCTGTGAAGTACAGCGCCAGCAAAAGTACAAATTTCTTGGAACTGGGT 780
DB 1210 ----- 1209
QY 781 AATGACCCAAATATATCTATGACACCATGATGCGCGGCAATGATGGCAGCCAGTTGGGA 840
DB 1210 ---GACCCAAATATATCTATGACACCATGATGCGCGGCAATGATGGCAGCCAGTTGGGA 1266
QY 841 AAGGATTATCATCCAGCTGAAGAGCTGTTGACGCCCATCGGATTTATTCAGAGCCAGC 900
DB 1267 AAGGATTATCATCCAGCTGAAGAGCTGTTGACGCCCATCGGATTTATTCAGAGCCAGC 1326
QY 901 TTAATATGCGCTTAATATTTGGGCGCGCGAGAGAAATGTCATGCGCTCTCTAGATGATTC 960
DB 1327 TTAATATGCGCTTAATATTTGGGCGCGCGAGAGAAATGTCATGCGCTCTCTAGATGATTC 1386
QY 961 ATGAGGTGGCAGAGATGACAGTAGATGACGTTACCTGGGCAACATTTGCTACATGATGCG 1020
DB 1387 ATGAGGTGGCAGAGATGACAGTAGATGACGTTACCTGGGCAACATTTGCTACATGATGCG 1446
QY 1021 CGGGTGGTCAAGGTGATGACCTTCTGAAAACCTGCGCTGTAGACACACTCTCTGACGAG 1080
DB 1447 CGGGTGGTCAAGGTGATGACCTTCTGAAAACCTGCGCTGTAGACACACTCTCTGACGAG 1506
QY 1081 ATTAGGAAAATTCAGAAAGTGTATATATACATACCTCCAGGAAAGAAATTTGGCTTGA 1140
DB 1507 ATTAGGAAAATTCAGAAAGTGTATATATACATACCTCCAGGAAAGAAATTTGGCTTGA 1566
QY 1141 GGTGTGGTGAACAACCTGAGTGAAGGCAAAACATCTATCCGATTCCTATGCTGACAGA 1200
DB 1567 GGTGTGGTGAACAACCTGAGTGAAGGCAAAACATCTATCCGATTCCTATGCTGACAGA 1626

```

```

QY 1201 TTCTTATGTTGAACATTTAGGAATGCTGGCCCAATCAAGGCAATTGATGTCGATACGG 1260
DB 1527 TTCTTATGTTGAACATTTAGGAATGCTGGCCCAATCAAGGCAATTGATGTCGATACGG 1686
QY 1261 CACTCATTTTGTGACCATGATACATCACTGCTGGAACAGAAATTTTAAACCATTAACA 1320
DB 1687 CACTCATTTTGTGACCATGATACATCACTGCTGGAACAGAAATTTTAAACCATTAACA 1746
QY 1321 GACTACTGCGCTCTCTCTCTCAAGCGCCTGATCGGCCCAAGTCTTGCTGTCAT 1380
DB 1747 GACTACTGCGCTCTCTCTCTCAAGCGCCTGATCGGCCCAAGTCTTGCTGTCAT 1806
QY 1381 GTGCTGCGGCTCCAGCGGAAACCAAGGCTGCGCGGATGATCCGGAACAACTTAAGAT 1440
DB 1807 GTGCTGCGGCTCCAGCGGAAACCAAGGCTGCGCGGATGATCCGGAACAACTTAAGAT 1866
QY 1441 TATGCTCACTGCAAAACCAACCAACCACTAGCTTGTGGGTTCATTAACATTTT 1500
DB 1867 TATGCTCACTGCAAAACCAACCAACCACTAGCTTGTGGGTTCATTAACATTTT 1926
QY 1501 ATCATCAACTTGCATGATCAAGAAAGAAATCAAGCTGCTGGAATCTCAAGACAG 1560
DB 1927 ATCATCAACTTGCATGATCAAGAAAGAAATCAAGCTGCTGGAATCTCAAGACAG 1986
QY 1561 CTGGTTGACCAAGTACCTGCTGACGCGCTATGGGAGGAGGCTTAAAGTCAAGTCAAGT 1620
DB 1987 CTGGTTGACCAAGTACCTGCTGACGCGCTATGGGAGGAGGCTTAAAGTCAAGTCAAGT 2046
QY 1621 CAACTGAATGGCCAGCCCTTATGATGATGAGAGACGCGGACCTCCCAAGATTAAGCCC 1680
DB 2047 CAACTGAATGGCCAGCCCTTATGATGATGAGAGACGCGGACCTCCCAAGATTAAGCCC 2106
QY 1681 CGCCCGCTTGGCGCGCGCGGACATTTGTCATCCCTCACTACACCAATGGCTTTTATGTG 1740
DB 2107 CGCCCGCTTGGCGCGCGCGGACATTTGTCATCCCTCACTACACCAATGGCTTTTATGTG 2166
QY 1741 GTCAAGATGTCATGCTTGGCTGCGCTACCGATTA 1779
DB 2167 GTCAAGATGTCATGCTTGGCTGCGCTACCGATTA 2205

RESULT 12
AAD29203
ID AAD29203 standard; cDNA; 1674 BP.
XX
AC AAD29203;
XX
DT 07-MAY-2002 (first entry)
XX
DE Human heparanase-2 splice variant cDNA.
XX
KW Human; heparanase-2; Hep-2; wound healing; angiogenesis; restenosis;
XX atherosclerosis; neurodegenerative disease; inflammation; prostanine;
XX viral infection; autoimmune lesion; renal failure; pancreatic cancer;
XX dystrophic muscular disease; heart disease; gene therapy; enzyme; ss.
OS Homo sapiens.
XX
PN MO200204645-A2.
XX
PD 17-JAN-2002.
XX
PF 12-JUL-2001; 2001WO-EP008094.
XX
PR 12-JUL-2000; 2000EP-00202442.
XX
PA (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
PI David G, Duerf J;
XX
DR WPI, 2002-171719/22.
XX
PT Heparanase-2 polypeptides and polymucleotides, useful for useful in wound

```

healing, angiogenesis, and for treating restenosis, atherosclerosis, inflammation, neurodegenerative diseases, and viral infections.

Disclosure; Page 40-41; 54p; English.

The invention relates to human heparanase-2 (Hep-2) polypeptides and polynucleotides. Heparanase-2 protein is useful in wound healing, angiogenesis, restenosis, atherosclerosis, neurodegenerative diseases, inflammation and viral infections, as well as in neutralising plasma heparin as a potential replacement of protamine. Antiheparanase-2 antibodies may be used for immunodetection and diagnosis of micrometastases, autoimmune lesions, renal failure in biopsy specimens, plasma samples and body fluids. Molecules which can agonise or antagonise heparanase-2 catalytic activity may also be used as a medicament. Polymorphisms in the polynucleotide sequence are useful in the identification of individuals having a predisposition to acquire diseases resulting from an increased or decreased expression of their activity. Such molecules can be used to treat pancreatic cancer, dystrophic muscular diseases and/or heart diseases. Polynucleotides of the invention are used in gene therapy. The present sequence is human heparanase-2 splice variant cDNA

Sequence 1674 BP; 419 A; 449 C; 421 G; 384 T; 0 U; 1 Other;

Query Match 79.3%; Score 1411; DB 6; Length 1674;

Best Local Similarity 89.8%; Pred. No. 0;

Matches 1598; Conservative 1; Mismatches 6; Indels 174; Gaps 1;

```

QY 1  ATGAGGAGTCTTGTGACCTCCCTGAGAGCAGTCCCTCAGAGCACTCCCGCCCGCCCGC 60
DB 56  ATGAGGAGTCTTGTGACCTCCCTGAGAGCAGTCCCTCAGAGCACTCCCGCCCGCCCGC 115
QY 61  TGCCTAGCCCGGAGGAGCTCTACTTGTGCTGTGCTCATCTCTCTCTCTCTCTCTCT 120
DB 116  TGCCTAGCCCGGAGGAGCTCTACTTGTGCTGTGCTCATCTCTCTCTCTCTCTCTCT 175
QY 121  GCTGGAGACAGAGACCTTGTGCTGTGACAGAGCTGCGAGTTTGAAGAAAAGCCTG 180
DB 176  GCTGGAGACAGAGACCTTGTGCTGTGACAGAGCTGCGAGTTTGAAGAAAAGCCTG 235
QY 181  ATTCTACTTGATGTGACACCAAGAACCCAGTCAAGACAGTCAATGAGAACTTCTCTCT 240
DB 236  ATTCTACTTGATGTGACACCAAGAACCCAGTCAAGACAGTCAATGAGAACTTCTCTCT 295
QY 241  CTGCACTGATGATCGTCAATCATGATGCTGCTGATTTCTTCTTCTTCTTCTTCTTCT 300
DB 296  CTGCACTGATGATCGTCAATCATGATGCTGCTGATTTCTTCTTCTTCTTCTTCTTCT 355
QY 301  TTGGTGAACCTGGCCCGGAGGAGCTTTCGCGCGCTTCTGCGCTTGGGGGGCAAAAGAAC 360
DB 356  TTGGTGAACCTGGCCCGGAGGAGCTTTCGCGCGCTTCTGCGCTTGGGGGGCAAAAGAAC 415
QY 361  GACTTCTTGCAGTTCAGAACTGAGAAACCCGCGAAAGCCGCGGAGCCCGGAGCCCG 420
DB 416  GACTTCTTGCAGTTCAGAACTGAGAAACCCGCGAAAGCCGCGGAGCCCGGAGCCCG 475
QY 421  GATTACTATCTCAAAAGATGAGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 476  GATTACTATCTCAAAAGATGAGATGATGATGATGATGATGATGATGATGATGATGAT 535
QY 481  CAGAAAGGCTGCAAGATGCTCCAGCACTGATGATGATGATGATGATGATGATGATGAT 540
DB 536  CAGAAAGGCTGCAAGATGCTCCAGCACTGATGATGATGATGATGATGATGATGATGAT 595
QY 541  GAGGCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 596  GAGGCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 655
QY 601  AATATTAACAGCAGGCTCTTGAACAACTTATTAATCTGATGATGATGATGATGATGAT 660
DB 656  AATATTAACAGGCTCTTGAACAACTTATTAATCTGATGATGATGATGATGATGATGAT 664
QY 661  CTGATATTTGCTTAATGCACTGGGCTGATATCCCAATTAATCTGGAACAGTTCTAGT 720

```

```

DB 665  ----- 664
QY 721  GCCCTGAGTCTGTGAAGTACAGCCGACAGAAAAGTACACATTTCTTGGAACTGGGT 780
DB 665  ----- 664
QY 781  AATGAGCCAAATACTATGAGACATGATGCGCCGAGTAAATGGACCCAGTTGGGA 840
DB 665  ---GAGCCAAATATACTATGAGACATGATGCGCCGAGTAAATGGACCCAGTTGGGA 721
QY 841  AAGATTAATCACTGAGTGAAGAGCTGTGACAGCCCATCCGATTTATTCAGAGCCAC 900
DB 722  AAGATTAATCACTGAGTGAAGAGCTGTGACAGCCCATCCGATTTATTCAGAGCCAC 781
QY 901  TTATATGCTCTTAATATTTGGGCGCGGAGAGAAATGTCATGCGCTCTTATGATGAT 960
DB 782  TTATATGCTCTTAATATTTGGGCGCGGAGAGAAATGTCATGCGCTCTTATGATGAT 841
QY 961  ATGAGAGTGGCAGAGAGTACAGTATGAGTACCTGCGCAACATTTGATCAATTGATG 1020
DB 842  ATGAGAGTGGCAGAGAGTACAGTATGAGTACCTGCGCAACATTTGATCAATTGATG 901
QY 1021  CGGCTGCTCAAGTATGATGATCTCTGAAAACCTGCTGTATGACACATCTCTGACCG 1080
DB 902  CGGCTGCTCAAGTATGATGATCTCTGAAAACCTGCTGTATGACACATCTCTGACCG 961
QY 1081  ATTAGAAAATTGAGAAAGTGTAAATACATACCTCCAGAGAAAGAAATTTGGCTTGA 1140
DB 962  ATTAGAAAATTGAGAAAGTGTAAATACATACCTCCAGAGAAAGAAATTTGGCTTGA 1021
QY 1141  GGTGTGTGACCACTTCACTGAGGAGCAAAACATCTATCTGATGCTGACGA 1200
DB 1022  GGTGTGTGACCACTTCACTGAGGAGCAAAACATCTATCTGATGCTGACGA 1081
QY 1201  TTCTTATGATGAACCTTTAGGAATGCTGGCCCAATCAGGAGCTGATGCTGATACGG 1260
DB 1082  TTCTTATGATGAACCTTTAGGAATGCTGGCCCAATCAGGAGCTGATGCTGATACGG 1141
QY 1261  CACTCATTTTGTGACATGAGTACATCACTCTGTGAGCCAGAAATTTTAAACCATTA 1320
DB 1142  CACTCATTTTGTGACATGAGTACATCACTCTGTGAGCCAGAAATTTTAAACCATTA 1201
QY 1321  GACTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380
DB 1202  GACTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1261
QY 1381  GTGGCTGGGCTCAGAGGAGGAGCCAGGCTGAGCGGAGTATCCGGGACAACTTAAGAT 1440
DB 1262  GTGGCTGGGCTCAGAGGAGGAGCCAGGCTGAGCGGAGTATCCGGGACAACTTAAGAT 1321
QY 1441  TATGCTACATGACAAACCAACCAACCACTTATGCTGTGAGTCACTTTT 1500
DB 1322  TATGCTACATGACAAACCAACCAACCACTTATGCTGTGAGTCACTTTT 1381
QY 1501  ATCATCACTTGTATGATCAAGAAAGAAATCACTGCTGAGCTCTCAGAGCAAG 1560
DB 1382  ATCATCACTTGTATGATCAAGAAAGAAATCACTGCTGAGCTCTCAGAGCAAG 1441
QY 1561  CTGGTTACAGATGCTGCTGCAAGCCCTATGGGACAGAGGCTTAAAGTCCAAAGT 1620
DB 1442  CTGGTTACAGATGCTGCTGCAAGCCCTATGGGACAGAGGCTTAAAGTCCAAAGT 1501
QY 1621  CAATGATAGGAGCCCTTATGATGATGAGAGAGAGCCCTCCCAATTTGAAGCC 1680
DB 1502  CAATGATAGGAGCCCTTATGATGATGAGAGAGAGCCCTCCCAATTTGAAGCC 1561
QY 1681  CGCCCTTTGGGCGGAGGAGATTTGATGATCTCTCAGTCACTATGAGCTTTATG 1740
DB 1562  CGCCCTTTGGGCGGAGGAGATTTGATGATCTCTCAGTCACTATGAGCTTTATG 1621
QY 1741  GTCAAGATGTCATGCTTTGGCTGCTGAGCTGAGTAA 1779

```

Db 1622 GTCAGATGTCATGCTTTGGCCTGCCGCTACCGATAA 1660

RESULT 13  
AAH28347  
ID AAH28347 standard; cDNA; 3943 BP.

DT 05-SEP-2001 (first entry)

Nucleotide sequence of human heparanase-like polypeptide.

KW Human; asparaginase-like polypeptide; gene therapy; cancer; angiogenesis;  
KW trauma; autoimmune disease; skin disease; cardiovascular disease;  
KW nervous system disease; inflammation; arthritis; genitalia;  
KW male fertility; erectile dysfunction; ss.

OS Homo sapiens

FH	Key	Location/Qualifiers
FT	CDS	1. .1479

PN WO200148161-A2.

PD 05-JUL-2001

PF 18-DEC-2000; 2000WO-EP012909.

PR 23-DEC-1999; 99EP-00125831.

PA (SCHD ) SCHERING AG

PI Siemeister G, Weiss B,

DR WPI; 2001-418259/44

DR P-PSDB; AAB84664.

PT Human Heparanase-like polynucleotide encoding polypeptides useful for  
PT modulating expression of the polypeptide and for treating cancer, cancer  
PT metastasis, aberrant angiogenesis by gene therapy technique.

PS Claim 1; Page 27-29; 30pp; English.

The present sequence encodes a human heparanase-like polypeptide. Heparanase-like polynucleotides are useful as a source of probes, primers and antisense molecules, and in gene therapy. Heparanase-like polynucleotides and polypeptides are useful for treating several disorders e.g., cancer, cancer metastasis. The oligonucleotides are also useful as diagnostic markers for the diagnosis of disorder such as cancer, cancer metastasis and aberrant angiogenesis. They may also act as diagnostic markers for diagnosis of disorder such as cancer, cancer metastasis and aberrant angiogenesis. The heparanase polypeptides and polynucleotides are also useful for treating trauma, autoimmune diseases, skin diseases, cardiovascular diseases, nervous system diseases, and inflammation including arthritis. Since the polynucleotide is preferentially expressed in male genitalia, modulation of its expression and/or activity may be used for medical intervention in male genitalia function that is male fertility control, erectile dysfunction

SQ Sequence 3943 BP; 1089 A; 921 C; 996 G; 934 T; 0 U; 3 Other;

Query Match	72.5%;	Score 1289.8;	DB 5;	Length 3943;
Best Local Similarity	89.2%;	Pred: NO. 0;		
Matches 1475;	Conservative	1;	Mismatches 3;	Indels 174;
				Gaps 1.

**Oy**      127 GACAGGAGACCCTTGCCCTGTAGA CAGACTGCAGGTTTGAAGGAAAAAGACCCTTAATTCTA 180  
|||||  
|||||  
**Db**      1 GACAGGAGACCCTTGCCCTGTAGA CAGACTGCAGGTTTGAAGGAAAAAGACCCTTAATTCTA 60

Oy	187	CTTGAATGTGAGCAACCAAGAACCCAGATCAGAGCATCATAGAACTCTCTCTCTGCAAG	246
Db	61	CTTGAATGTGAGCAACCAAGAACCCAGATCAGAGCATCATAGAACTCTCTCTCTGCAAG	120
Oy	247	CTGGATCCGTCATCATTCATCATATGATGCTGGCTCGATTTCTTAAGCTCCAGCGCTTGGTG	306
Db	121	CTGGATCCGTCATCATTCATCATATGATGCTGGCTCGATTTCTTAAGCTCCAGCGCTTGGTG	180
Oy	307	ACCGTGGCCCGGGGACTTTGGCCGCGCTTTCTGGCGCTTCGGGGGCAAAAGACCGACTTC	366
Db	181	ACCGTGGCCCGGGGACTTTGGCCGCGCTTTCTGGCGCTTCGGGGGCAAAAGACCGACTTC	240
Oy	367	CTGCAGTTCAGAACTGTAGAAACCCTGGCGAAAAGCCCGGGGGGCCCGGGCCGGATTAAC	426
Db	241	CTGCAGTTCAGAACTGTAGAAACCCTGGCGAAAAGCCCGGGGGGCCCGGGCCGGATTAAC	300
Oy	427	TATCTCAAAAACATATAGATGATGACATGTTCGAAGTATGTGCGCTTATATTAACAGAAA	486
Db	301	TATCTCAAAAACATATAGATGATGACATGTTCGAAGTATGTGCGCTTATATTAACAGAAA	360
Oy	487	GGCTGCAAGATTGGCCCGACACCTGTATGTATGTCTGTGCTCCAAAGGAGAGGACGCT	546
Db	361	GGCTGCAAGATTGGCCCGACACCTGTATGTATGTCTGTGCTCCAAAGGAGAGGACGCT	420
Oy	547	CAGATGCATCTGGTCTCTTAAAGGACCAATTCCTCAATCTTACAGTATCTCATATTA	606
Db	421	CAGATGCATCTGGTCTCTTAAAGGACCAATTCCTCAATCTTACAGTATCTCATATTA	480
Oy	607	ACAGCCAGGTCTCTAGCAAACTTTATPACTTTGCTGATTTCTGTGACTCCACCTGATA	666
Db	481	ACA-----	483
Oy	667	TTTGCTCTAAATGCACGTGCGTGTATTCCAATACTCTTGAAACAGTTCTAGTCCCTG	726
Db	484	-----	483
Oy	727	AGTCTGTGAAGTACAGCGCCAGCAAAAAGTACAACATTTCTTGAGAACTGGGTAAATGAG	786
Db	484	-----GAG	486
Oy	787	CCAATATACATTCGGAACAATGCATGGCCGGGACAGTAAATGGACGCGAGTTGGAAAAGAT	846
Db	487	CCAATATACATTCGGAACAATGCATGGCCGGGACAGTAAATGGACGCGAGTTGGAAAAGAT	546
Oy	847	TACATCCAGCTGGAAGAGCTTGTGACGCCATCCGGAATTTATTTTCCAGAGCCAGCTTATAT	906
Db	547	TACATCCAGCTGGAAGAGCTTGTGACGCCATCCGGAATTTATTTTCCAGAGCCAGCTTATAT	606
Oy	907	GGCCCTAATATTTGGGGCGGCGGAGGAAGATTCATGCGCCCTCTAGATGTGATTCATGAAG	966
Db	607	GGCCCTAATATTTGGGGCGGCGGAGGAAGATTCATGCGCCCTCTAGATGTGATTCATGAAG	666
Oy	967	GTGGCAGGAATACATGATGATGACAGTTTACTGTGCAACATTTGCTATATGATGCTCGGGTG	1026
Db	667	GTGGCAGGAATACATGATGATGACAGTTTACTGTGCAACATTTGCTATATGATGCTCGGGTG	726
Oy	1027	GTCAAGGTGATGAGACTTCCTGAAAACCTGCGCTGTGAACAACACTCTCTGACCAAGATTAGG	1086
Db	727	GTCAAGGTGATGAGACTTCCTGAAAACCTGCGCTGTGAACAACACTCTCTGACCAAGATTAGG	786
Oy	1087	AAATATCAGAAATGTGTTATATACATPACATCTCCAGAAAAGAAATTTGGCTTGAAGGTGTG	1146
Db	787	AAATATCAGAAATGTGTTATATACATPACATCTCCAGAAAAGAAATTTGGCTTGAAGGTGTG	846
Oy	1147	GTGACCACTCAGCTGAGGAGGACAAACAATCTATCCGATTCCTATNGCTGACAGATTCTTA	1206
Db	847	GTGACCACTCAGCTGAGGAGGACAAACAATCTATCCGATTCCTATNGCTGACAGATTCTTA	906
Oy	1207	TGTTTGAACACTTTAGGAATGCTGGCCCATCAGGGCAATTGATGTCTGTATACGGCACTCA	1266
Db	907	TGTTTGAACACTTTAGGAATGCTGGCCCATCAGGGCAATTGATGTCTGTATACGGCACTCA	966



```

Db      534 ATCCGATTTATTCACAGCCAGCTTATATGGCCCTAAATTGGGGGGCCGAGAGAAAGAT 593
Qy      937 GTATACGCCCTCTCTAGATGAGATTCATTAAGTGTGACAGAAATGACATGATGCACTTACC 996
Db      594 GTATACGCCCTCTCTAGATGAGATTCATTAAGTGTGACAGAAATGACATGATGCACTTACC 653
Qy      997 TGGCAACATTTCTACATTTATGAGCCGGGTGTCAAGGTGATGATCTCCGAAATCTCGC 1056
Db      654 TGGCAACATTTCTACATTTATGAGCCGGGTGTCAAGGTGATGATCTCCGAAATCTCGC 713
Qy      1057 CTGTGAACACACTCTCTGACACAGATTTAGAAAATTCAGAAAGTGGTTAATACATCACT 1116
Db      714 CTGTGAACACACTCTCTGACACAGATTTAGAAAATTCAGAAAGTGGTTAATACATCACT 773
Qy      1117 CCAAGAAAGAGATTTGGCTTGAAGGTGTGACCACTCAGCTGGAGGCAAAACAAAT 1176
Db      774 CCAAGAAAGAGATTTGGCTTGAAGGTGTGACCACTCAGCTGGAGGCAAAACAAAT 833
Qy      1177 CTATCCGATTTCTATGCTGACAGATTTCTTATGGTTGAACAATTAGGAATGCTGGCCAAAT 1236
Db      834 CTATCCGATTTCTATGCTGACAGATTTCTTATGGTTGAACAATTAGGAATGCTGGCCAAAT 893
Qy      1237 CAGGGCATTTATGCTGATACCGGCACTCATTTTGTGACATGATGATCAATCACTCGTG 1296
Db      894 CAGGGCATTTATGCTGATACCGGCACTCATTTTGTGACATGATGATCAATCACTCGTG 953
Qy      1297 GACCAAGATTTTAAACCATTAACAGACATCTGCTCTCTCTCTCAACAAGCCCTGATC 1356
Db      954 GACCAAGATTTTAAACCATTAACAGACATCTGCTCTCTCTCTCAACAAGCCCTGATC 1013
Qy      1357 GGGCCCAAACTTTGGCTGTGATGATGCTGGGCTCCAGGGAAGCAAGGCTGGCCCA 1416
Db      1014 GGGCCCAAACTTTGGCTGTGATGATGCTGGGCTCCAGGGAAGCAAGGCTGGCCCA 1073
Qy      1417 GTGATCCGGGACAAACTAAGATTTATGCTCACTGCAAAACCAACCAACAACAATC 1476
Db      1074 GTGATCCGGGACAAACTAAGATTTATGCTCACTGCAAAACCAACAACAACAACAATC 1133
Qy      1477 GTTGTGGGTGCTATTAACCTTTTATCATCACTTGCATGATCAAGAAAGAAATCAAG 1536
Db      1134 GTTGTGGGTGCTATTAACCTTTTATCATCACTTGCATGATCAAGAAAGAAATCAAG 1193
Qy      1537 CTGAGTGGAGCTCTCAGAGACAAAGCTGTTCAACAGTACCTGCTGAGGCCATATGGGACG 1596
Db      1194 CTGAGTGGAGCTCTCAGAGACAAAGCTGTTCAACAGTACCTGCTGAGGCCATATGGGACG 1253
Qy      1597 GAGGGCTTAAGTTCAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1656
Db      1254 GAGGGCTTAAGTTCAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1313
Qy      1657 GAGGACCTCCCAATTTGAAGCCCCCGCCCTTCCGGGCGGCGGACATTTGGTCACTCCCT 1716
Db      1314 GAGGACCTCCCAATTTGAAGCCCCCGCCCTTCCGGGCGGCGGACATTTGGTCACTCCCT 1373
Qy      1717 CCAAGTCAACATGAGGCTTTTATGCTGATCAAGATGATGATGATGATGATGATGATGATG 1776
Db      1374 CCAAGTCAACATGAGGCTTTTATGCTGATCAAGATGATGATGATGATGATGATGATG 1433
Qy      1777 TAA 1779
Db      1434 TAA 1436

```

```

RESULT 15
ID      AAA91099
AC      AAA91099;
XX      20-APR-2001 (first entry)
DE      Human heparanase, hmbpl pns form, coding sequence.

```

```

XX      Heparanase; hmbpl; wound healing; angiogenesis; restenosis; Scurvy;
KM      atherosclerosis; inflammation; pulmonary disease; Alzheimer's disease;
KM      neurodegenerative disease; Creutzfeldt-Jakob disease; viral infection;
KM      gene therapy; human; ds.
OS      Homo sapiens.
FH      Key      Location/Qualifiers
FT      CDS      26..1468
FT      /tag=      a
FT      /product= "heparanase"
PN      WO200100643-A2.
PD      04-JAN-2001.
PF      19-JUN-2000; 2000MO-IL000358.
PR      25-JUN-1999; 99US-0140801P.
PA      (INSI-) INSIGHT STRATEGY & MARKETING LTD.
XX      Pecker I, Michal I, Itzhaki H;
XX      WPI; 2001-137930/14.
DR      P-PSDB; AAY97634.
XX      New polynucleotides and polypeptides that are distantly homologous to
PT      heparanase, useful in wound healing, as well as in gene therapy protocols
PT      for angiogenesis, restenosis, atherosclerosis, or inflammation.
XX      Claim 3; Page 62; 67P; English.
XX      This sequence encodes a heparanase of the invention. The heparanase DNA
CC      and protein sequences are useful in wound healing, angiogenesis,
CC      restenosis, atherosclerosis, inflammation, pulmonary diseases,
CC      neurodegenerative diseases (such as Scurvy, Alzheimer's disease, and
CC      Creutzfeldt-Jakob disease) or viral infections. The heparanase coding
CC      sequence is particularly useful in gene therapy.
XX      Sequence 1724 BP; 423 A; 481 C; 422 G; 398 T; 0 U; 0 Other;
SQ

```

```

Query Match      61.6%; Score 1095.4; DB 4; Length 1724;
Best Local Similarity 81.1%; Pred. No. 0;
Matches 1442; Conservative 0; Mismatches 1; Indels 336; Gaps 1;
Qy      1 ATGAGGGTGTCTTTGTGCTTCCCTGAAGCCATGCGCTCCAGCACTCCGCCCCCGCG 60
Db      26 ATGAGGGTGTCTTTGTGCTTCCCTGAAGCCATGCGCTCCAGCACTCCGCCCCCGCG 85
Qy      61 TGCCTAGCCCCGGGGGCTCTTACTTGGCTGTGCTGCTCCATCTCTCCCTTCTCCAG 120
Db      86 TGCCTAGCCCCGGGGGCTCTTACTTGGCTGTGCTGCTCCATCTCTCCCTTCTCCAG 145
Qy      121 GCTGAGACAGAGAACCTTTCGCTGAGACAGAGCTGAGGTTGAAGAAAGAACCTCTG 180
Db      146 GCTGAGACAGAGAACCTTTCGCTGAGACAGAGCTGAGGTTGAAGAAAGAACCTCTG 205
Qy      181 ATTCTACTTGTATGTGAGACCAAGAACCCAGTCAAGACAGTCAATGAGAACTTCTCTCT 240
Db      206 ATTCTACTTGTATGTGAGACCAAGAACCCAGTCAAGACAGTCAATGAGAACTTCTCTCT 265
Qy      241 CTGACAGCTGATTCGCTGCATCATGATGAGTGGCTGGATTTCTTAAGCTCCAGAGCG 300
Db      266 CTGACAGCTGATTCGCTGCATCATGATGAGTGGCTGGATTTCTTAAGCTCCAGAGCG 325
Qy      301 TTGATGACCTTGCGCCGGGGAATTGTGCGCCCTTTCTGCGCTTGGGGGCAAAAGAGCC 360
Db      326 TTGATGACCTTGCGCCGGGGAATTGTGCGCCCTTTCTGCGCTTGGGGGCAAAAGAGCC 385
Qy      361 GACTTCCGACGTTCCAGAACCTGAGAACCCGGCGAAAGCCGGCGGGGCGCGGCGCG 420

```











Db 601 ATGAGGAGTCTTTGGCTTCCCTGAGGCAATGCCCTCAGCAATCCCGCCCGCCG 660  
 Qy 61 TGGCTAGAGCCGGGGGCTCTACTAGGCTCTGTTGCTCATCTCTCCCTTCTCCAG 120  
 Db 661 TGGCTAGAGCCGGGGGCTCTACTAGGCTCTGTTGCTCATCTCTCCCTTCTCCAG 720  
 Qy 121 GCTGAGACAGAGACCTTGGCTGTAGACAGAGCTGACGGTTTGAAGAAAAGACCTG 180  
 Db 721 GCTGAGACAGAGACCTTGGCTGTAGACAGAGCTGACGGTTTGAAGAAAAGACCTG 780  
 Qy 181 ATTCTACTTGTAGTGGACACCAAGACAGTCAAGGACAGTCAATGAGAACTTCTCT 240  
 Db 781 ATTCTACTTGTAGTGGACACCAAGACAGTCAAGGACAGTCAATGAGAACTTCTCT 840  
 Qy 241 CTGACAGCTGGATCCGTCAATCATTCATGATGGCTGGCTGATTTCTTAAGCTCAAGGC 300  
 Db 841 CTGACAGCTGGATCCGTCAATCATTCATGATGGCTGGCTGATTTCTTAAGCTCAAGGC 900  
 Qy 301 TTGGTACCTGGCCCGGGGACCTTTGGCCGCTTCTGCGCTTGGGGGCAAAAGAAC 360  
 Db 901 TTGGTACCTGGCCCGGGGACCTTTGGCCGCTTCTGCGCTTGGGGGCAAAAGAAC 960  
 Qy 361 GACTTCCTGAGTTCAGAACTGAGGAAACCCGCGCAAAAGCCGCGGGCCCGCGCCG 420  
 Db 961 GACTTCCTGAGTTCAGAACTGAGGAAACCCGCGCAAAAGCCGCGGGCCCGCGCCG 1020  
 Qy 421 GATTACTATTCAAAACTATGAGGATGACTTTGTCGAAGTATGTGCTTAAATAA 480  
 Db 1021 GATTACTATTCAAAACTATGAGGATGACTTTGTCGAAGTATGTGCTTAAATAA 1080  
 Qy 481 CAGAAAGGCTGCAAGATTGCGCAGCACTGATGTTATGCTGGTGTCCAAAGGAGAA 540  
 Db 1081 CAGAAAGGCTGCAAGATTGCGCAGCACTGATGTTATGCTGGTGTCCAAAGGAGAA 1140  
 Qy 541 GCAGCTCAGATGATCTGTTCTTTCTAAAGAGCAATTCCTCAATCTTAAGATCTC 600  
 Db 1141 GCAGCTCAGATGATCTGTTCTTTCTAAAGAGCAATTCCTCAATCTTAAGATCTC 1200  
 Qy 601 AATTAAACAGCAGGCTCTAGACAACTTAACTTATGCTGATGCTGTGACTCCAG 660  
 Db 1201 AATTAAACAGCAGGCTCTAGACAACTTAACTTATGCTGATGCTGTGACTCCAG 1260  
 Qy 661 CTGATATTTGCTCTAAATGCACTGCGTCTGATTCCTCAATCTCTGGAACAGTTCTAGT 720  
 Db 1261 CTGATATTTGCTCTAAATGCACTGCGTCTGATTCCTCAATCTCTGGAACAGTTCTAGT 1320  
 Qy 721 GCCCTGAGTCTGTGAGTACAGGCGCAGCAAAAGTAAACAATTTCTGGGAACTGGGT 780  
 Db 1321 GCCCTGAGTCTGTGAGTACAGGCGCAGCAAAAGTAAACAATTTCTGGGAACTGGGT 1380  
 Qy 781 AATGAGCCAAATACTATCGGACCATGATGCGCGCAGTAAATGCGACCCAGTTGGGA 840  
 Db 1381 AATGAGCCAAATACTATCGGACCATGATGCGCGCAGTAAATGCGACCCAGTTGGGA 1440  
 Qy 841 AAGGATTACATCCAGCTGAGAGAGCTTGTGACGCCATCCGGAATTTATCCAGAGCCAGC 900  
 Db 1441 AAGGATTACATCCAGCTGAGAGAGCTTGTGACGCCATCCGGAATTTATCCAGAGCCAGC 1500  
 Qy 901 TTATATGAGCCCTATATTTGGCGCGCGCAGGAAAGATGCTATGCGCTCTAGATGATTC 960  
 Db 1501 TTATATGAGCCCTATATTTGGCGCGCGCAGGAAAGATGCTATGCGCTCTAGATGATTC 1560  
 Qy 961 ATGAAAGTGGCAGAGAGTACAGTATGAGCACTTCTGGCAACATTTCTCATTTGATGGC 1020  
 Db 1561 ATGAAAGTGGCAGAGAGTACAGTATGAGCACTTCTGGCAACATTTCTCATTTGATGGC 1620  
 Qy 1021 CGGGTGGTCAAGTATGATGATCTTCTGTAAACTCGCTGTATGACACACTCTGTACAG 1080  
 Db 1621 CGGGTGGTCAAGTATGATGATCTTCTGTAAACTCGCTGTATGACACACTCTGTACAG 1680  
 Qy 1081 ATTAGAAATTCAGAAAGTGTATATACATCACTCCAGGAAAGAAATTTGGCTTGA 1140

Db 1681 ATTAGAAATTCAGAAAGTGTATATACATCACTCCAGGAAAGAAATTTGGCTTGA 1740  
 Qy 1141 GGTGTGTGACCACTCTGAGTGGAGGACAAACATCTATCCGATTCCTATGTGACAGA 1200  
 Db 1741 GGTGTGTGACCACTCTGAGTGGAGGACAAACATCTATCCGATTCCTATGTGACAGA 1800  
 Qy 1201 TTCTTATGTTGAACCTTATGAGATGTGGCCATCAGGGCATTTGATGCTGATACGG 1260  
 Db 1801 TTCTTATGTTGAACCTTATGAGATGTGGCCATCAGGGCATTTGATGCTGATACGG 1860  
 Qy 1261 CACTCATTTTATGACCAATGATACATCACTCTGTGAGACAGAAATTTAAACCATTAACA 1320  
 Db 1861 CACTCATTTTATGACCAATGATACATCACTCTGTGAGACAGAAATTTAAACCATTAACA 1920  
 Qy 1321 GACTACTGGCTCTCTCTCTCTCAAGCGCTGATATGGGCCCAAGCTTGGCTGTGAT 1380  
 Db 1921 GACTACTGGCTCTCTCTCTCTCAAGCGCTGATATGGGCCCAAGCTTGGCTGTGAT 1980  
 Qy 1381 GTGGCTGGGCTCCAGCGAAAGCCAGCGCTGGCCAGTATCCGGACAAACTPAAGATT 1440  
 Db 1981 GTGGCTGGGCTCCAGCGAAAGCCAGCGCTGGCCAGTATCCGGACAAACTPAAGATT 2040  
 Qy 1441 TATGCTCACTGCAACAAACCAACCAACCAACCACTAGCTGTGGGTCATTACACTTTT 1500  
 Db 2041 TATGCTCACTGCAACAAACCAACCAACCAACCACTAGCTGTGGGTCATTACACTTTT 2100  
 Qy 1501 ATCATCAACTGCATCGATCAAGAAAGAAATCAAGCTGGCTGGGACTCTCAGAGACAG 1560  
 Db 2101 ATCATCAACTGCATCGATCAAGAAAGAAATCAAGCTGGCTGGGACTCTCAGAGACAG 2160  
 Qy 1561 CTGGTTCACCACTACTCTGACGCCCTATGAGGAGAGAGGGCCCTAAGTCCAAATCAGTG 1620  
 Db 2161 CTGGTTCACCACTACTCTGACGCCCTATGAGGAGAGAGGGCCCTAAGTCCAAATCAGTG 2220  
 Qy 1621 CACTGAATGGCCAGCCCTTATGATGTGACAGCGGACCTCTCCAGAAATTAAGCC 1680  
 Db 2221 CACTGAATGGCCAGCCCTTATGATGTGACAGCGGACCTCTCCAGAAATTAAGCC 2280  
 Qy 1681 CGCCCTTGGGGCGGCGCGGACATGCTCATTCCTCCAGTACACATGGGCTTTTATGTG 1740  
 Db 2281 CGCCCTTGGGGCGGCGCGGACATGCTCATTCCTCCAGTACACATGGGCTTTTATGTG 2340  
 Qy 1741 GTCAAGAAATGCAATGCTTGGCGCTGCGCTACCGATTA 1779  
 Db 2341 GTCAAGAAATGCAATGCTTGGCGCTGCGCTACCGATTA 2379

RESULT 3  
 US-10-302-172-409  
 ; Sequence 409, Application US/10302172  
 ; Publication No. US20040053250A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tang, Y. Tom  
 ; APPLICANT: Xue, Aidong J.  
 ; APPLICANT: Drmanac, Radjoje T.  
 ; TITLE OF INVENTION: No. US20040053250A1 Arginine-rich Protein-like Nucleic Acids and  
 ; FILE REFERENCE: 803 JCNCF  
 ; CURRENT APPLICATION NUMBER: US/10/302,172  
 ; CURRENT FILING DATE: 2002-11-21  
 ; PRIOR APPLICATION NUMBER: US 10/225,251  
 ; PRIOR FILING DATE: 2002-08-20  
 ; PRIOR APPLICATION NUMBER: PCT US02/05095  
 ; PRIOR FILING DATE: 2002-03-05  
 ; PRIOR APPLICATION NUMBER: US 09/799,451  
 ; NUMBER OF SEQ ID NOS: 950  
 ; SOFTWARE: pc\_ft\_genes Version 2.0  
 ; SEQ ID NO 409  
 ; LENGTH: 1891  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:

NAME/KEY: CDS  
LOCATION: (20)...(1621)  
US-10-302-172-409

Query Match 79.7%; Score 1417.8; DB 13; Length 1891;  
Best Local Similarity 90.1%; Pred. No. 0;  
Matches 1603; Conservative 0; Mismatches 2; Indels 174; Gaps 1;

QY 1 ATGAGGCTCTTGTGCTTCCCTGAGGCAATGCCCTCCAGCACTCCGCGCCCGCG 60  
DB 20 ATGAGGCTCTTGTGCTTCCCTGAGGCAATGCCCTCCAGCACTCCGCGCCCGCG 79  
QY 61 TGGCTGCCCCGGGGGCTCTTACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
DB 80 TGGCTGCCCCGGGGGCTCTTACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 139  
QY 121 GCTGAGACAGAGACCTTCTCTGAGACAGAGCTGAGCTTGAAGAGAAAGACCTG 180  
DB 140 GCTGAGACAGAGACCTTCTCTGAGACAGAGCTGAGCTTGAAGAGAAAGACCTG 199  
QY 181 ATTCTACTGATGTGACACCAAGAACCAAGTCAAGACAGTCAATGAGAACTTCTCT 240  
DB 200 ATTCTACTGATGTGACACCAAGAACCAAGTCAAGACAGTCAATGAGAACTTCTCT 259  
QY 241 CTGAGCTGATCCGTCATCATTCATGATGCTGCTGCTGCTGCTGCTGCTGCTG 300  
DB 260 CTGAGCTGATCCGTCATCATTCATGATGCTGCTGCTGCTGCTGCTGCTGCTG 319  
QY 301 TTGCTGACCTGCGCCGGGACCTTTCGCTTCTGCTGCTTCTGCTGCTTCTGCTG 360  
DB 320 TTGCTGACCTGCGCCGGGACCTTTCGCTTCTGCTGCTTCTGCTGCTTCTGCTG 379  
QY 361 GAGCTCTGAGTTCAGAACCTGAGGAACCGGCGGAGAAAGCGGCGGCGCGCG 420  
DB 380 GAGCTCTGAGTTCAGAACCTGAGGAACCGGCGGAGAAAGCGGCGGCGCGCGCG 439  
QY 421 GATTACTATCTCAAAAATCTAGAGATGACATTTGCGAAATGATGCTTCTGATATA 480  
DB 440 GATTACTATCTCAAAAATCTAGAGATGACATTTGCGAAATGATGCTTCTGATATA 499  
QY 481 CAGAAAGCTGCAAGATTTGCGAGACCTGATGATGCTGCTGCTGCTGCTGCTG 540  
DB 500 CAGAAAGCTGCAAGATTTGCGAGACCTGATGATGCTGCTGCTGCTGCTGCTG 559  
QY 541 GAGCTCAGATGATGCTGCTTCTTCAAGAGCAATCTTCAATCTTCAAGTATCTC 600  
DB 560 GAGCTCAGATGATGCTGCTTCTTCAAGAGCAATCTTCAATCTTCAAGTATCTC 619  
QY 601 ATTCTTACAGCCAGTCTCTAGACAACTTATTAATTTGCTGATGCTGCACTCCAC 660  
DB 620 ATTCTTACAA----- 628  
QY 661 CTGATATTTCTCTAATGCACTGCGTCTTATCTCCAAATACCTGGAACAGTTCTAGT 720  
DB 629 ----- 628  
QY 721 GCCCTGAGCTCTTGAAGTACAGCGCCAGCAAAAAGTAAACATTTCTTGGAACTGGGT 780  
DB 629 ----- 628  
QY 781 AATGAGCAATACTACTGCAACCATGATGCGGCGAGTAAATGCAAGCAGTTGGGA 840  
DB 629 ---GAGCAATACTACTGCAACCATGATGCGGCGAGTAAATGCAAGCAGTTGGGA 865  
QY 841 AAGGATTAATCCAGCTGAGAGAGCTGCTGCAAGCCCATCCGAGTTTATTCAGAGCCAGC 900  
DB 686 AAGGATTAATCCAGCTGAGAGAGCTGCTGCAAGCCCATCCGAGTTTATTCAGAGCCAGC 945  
QY 901 TTATATGCGCTATATTTGGGCGGCGAGAGAAAGATGCAATCCGCTCTAGATGATTC 960  
DB 746 TTATATGCGCTATATTTGGGCGGCGAGAGAAATGCAATCCGCTCTAGATGATTC 805  
QY 961 ATGAGGCTGCGAGAGATGACATGATGACATTAAGTGAACATTTGCTCATTTGATGC 1020

DB 806 ATGAGGCTGCGAGAGATGACATGATGACATTAAGTGAACATTTGCTCATTTGATGC 865  
QY 1021 CGGCTGCTCAAGGTGATGACATTTCTGAAAACTGCTGTTAGACACACTCTCTGACCA 1080  
DB 866 CGGCTGCTCAAGGTGATGACATTTCTGAAAACTGCTGTTAGACACACTCTCTGACCA 925  
QY 1081 ATTGAGAAATTTGAAAGTGTCTTATTAATCAATCACTCCAGAGAAAGAAATTGGCTTGA 1140  
DB 926 ATTGAGAAATTTGAAAGTGTCTTATTAATCAATCACTCCAGAGAAAGAAATTGGCTTGA 985  
QY 1141 GGTGTGTGACCACTGAGCTGAGGACAAACATCTATCCGATTCATGCTGACGA 1200  
DB 986 GGTGTGTGACCACTGAGCTGAGGACAAACATCTATCCGATTCATGCTGACGA 1045  
QY 1201 TTCTTATGTTGAACATTTAGAAATGCTGCGCAATCAAGGCAATTAATGCTGATACGG 1260  
DB 1046 TTCTTATGTTGAACATTTAGAAATGCTGCGCAATCAAGGCAATTAATGCTGATACGG 1105  
QY 1261 CACTCATTTTGTGACATGATTAATCAATCACTGATGAGACAGAAATTTAAACCATTTACA 1320  
DB 1106 CACTCATTTTGTGACATGATTAATCAATCACTGATGAGACAGAAATTTAAACCATTTACA 1165  
QY 1321 GACTTACTGCTCTCTCTCTCTCAAGCGCTGATGCGCCCAAGTCTTGGCTGTGCAT 1380  
DB 1166 GACTTACTGCTCTCTCTCTCTCAAGCGCTGATGCGCCCAAGTCTTGGCTGTGCAT 1225  
QY 1381 GTGCTGCGGCTCCAGGAGAAAGCAAGCGCTGCGAGATTCGGGGAACAATTAAGAT 1440  
DB 1226 GTGCTGCGGCTCCAGGAGAAAGCAAGCGCTGCGAGATTCGGGGAACAATTAAGAT 1285  
QY 1441 TATGCTCATGCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1500  
DB 1286 TATGCTCATGCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1345  
QY 1501 ATCTCAACTGCTGATCAAGAAAGAAATCAAGTGTGGGACTCTCAGAGACAAG 1560  
DB 1346 ATCTCAACTGCTGATCAAGAAAGAAATCAAGTGTGGGACTCTCAGAGACAAG 1405  
QY 1561 CTGCTTCAACAGTACTGCTGACAGCCCTATGAGGAGAGGCGCTTAAATCCAGTCA 1620  
DB 1406 CTGCTTCAACAGTACTGCTGACAGCCCTATGAGGAGAGGCGCTTAAATCCAGTCA 1465  
QY 1621 CACTGATGAGCCAGCCCTTATGATGATGAGAGAGAGGCGCTTAAATCCAGTCA 1680  
DB 1466 CACTGATGAGCCAGCCCTTATGATGATGAGAGAGAGGCGCTTAAATCCAGTCA 1525  
QY 1681 CGCCCGCTTCCGCGCGGCGGAGCAATGATCAATCCCTCAAGTCAACATGAGGCTTTATGTG 1740  
DB 1526 CGCCCGCTTCCGCGCGGCGGAGCAATGATCAATCCCTCAAGTCAACATGAGGCTTTATGTG 1585  
QY 1741 GTCAAGATGATGATGCTTGGCTGCGGCTACCGATTA 1779  
DB 1586 GTCAAGATGATGATGCTTGGCTGCGGCTACCGATTA 1624

RESULT 4  
US-09-836-461-1  
Sequence 1, Application US/09836461  
Patent No. US20020064853A1  
GENERAL INFORMATION:  
APPLICANT: Bienkowski, Michael J.  
APPLICANT: Heinrichson, Robert L.  
TITLE OF INVENTION: Heparanase II, A No. US20020064853A1 Human Heparanase Paralog  
FILE REFERENCE: heparanase II  
CURRENT APPLICATION NUMBER: US/09/836,461  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 2326  
TYPE: DNA  
ORGANISM: Homo sapiens

US-09-836-461-1

Query Match 79.7%; Score 1417.8; DB 9; Length 2326;  
 Best Local Similarity 90.1%; Pred. No. 0;  
 Matches 1603; Conservative 0; Mismatches 2; Indels 174; Gaps 1;

QY 1 ATGAGGCTGCTTTGCTGCTTCCCTGAAAGCATGCTCCAGCAACCTCCGCCCCCCCCG 60  
 DB ATGAGGCTGCTTTGCTGCTTCCCTGAAAGCATGCTCCAGCAACCTCCGCCCCCCCCG 84  
 QY 61 TGGCTAGCCCCGGGGGCTCTCTAATTGGCTCTGTGCTCAATCTCCCTTCTCCAG 120  
 DB TGGCTAGCCCCGGGGGCTCTCTAATTGGCTCTGTGCTCAATCTCCCTTCTCCAG 144  
 QY 121 GCTGAGACAGAGACCTTGGCTGTAGACAGAGCTGCAAGTTGAAGAAAGACCTG 180  
 DB GCTGAGACAGAGACCTTGGCTGTAGACAGAGCTGCAAGTTGAAGAAAGACCTG 204  
 QY 181 ATTCTAATTGATGTAGACACCAAGAACCCAGTGAAGACAGTCAATGAGACTCTCT 240  
 DB ATTCTAATTGATGTAGACACCAAGAACCCAGTGAAGACAGTCAATGAGACTCTCT 264  
 QY 241 CTGCACTGATCCGCTCCATCATTCATGATGCTGCTGATTTCTTAAGCTCCAGCC 300  
 DB CTGCACTGATCCGCTCCATCATTCATGATGCTGCTGATTTCTTAAGCTCCAGCC 324  
 QY 301 TTGGTGAACCTTGACCCGGGGACCTTTGGCCCTTTCTGGCTTCCGGGGCAAAAGAC 360  
 DB TTGGTGAACCTTGACCCGGGGACCTTTGGCCCTTTCTGGCTTCCGGGGCAAAAGAC 384  
 QY 325 TTGGTGAACCTTGACCCGGGGACCTTTGGCCCTTTCTGGCTTCCGGGGCAAAAGAC 444  
 DB GACTTCCTGAGTTCCAGAACCTGAGGAAACCCGGGGCAAAAGACCCGGGGCCG 444  
 QY 361 GACTTCCTGAGTTCCAGAACCTGAGGAAACCCGGGGCAAAAGACCCGGGGCCG 420  
 DB GACTTCCTGAGTTCCAGAACCTGAGGAAACCCGGGGCAAAAGACCCGGGGCCG 444  
 QY 421 GATTACTATCTCAAAAATACTAGAGATGATGTTGGAAGTATGTTGCTTAAGTAA 480  
 DB GATTACTATCTCAAAAATACTAGAGATGATGTTGGAAGTATGTTGCTTAAGTAA 504  
 QY 445 GATTACTATCTCAAAAATACTAGAGATGATGTTGGAAGTATGTTGCTTAAGTAA 504  
 DB GATTACTATCTCAAAAATACTAGAGATGATGTTGGAAGTATGTTGCTTAAGTAA 504  
 QY 481 CAGAAAGCTGCAAGATTTGCGCAACCTGATGTTAAGTGTGCTCCAAAGGAGAA 540  
 DB CAGAAAGCTGCAAGATTTGCGCAACCTGATGTTAAGTGTGCTCCAAAGGAGAA 564  
 QY 505 CAGAAAGCTGCAAGATTTGCGCAACCTGATGTTAAGTGTGCTCCAAAGGAGAA 564  
 DB CAGAAAGCTGCAAGATTTGCGCAACCTGATGTTAAGTGTGCTCCAAAGGAGAA 564  
 QY 541 GCAAGCTCAATGATCTGCTTTCTTAAGAGAAATCTCCAAATCTTAAGTATCTC 600  
 DB GCAAGCTCAATGATCTGCTTTCTTAAGAGAAATCTCCAAATCTTAAGTATCTC 624  
 QY 565 GCAAGCTCAATGATCTGCTTTCTTAAGAGAAATCTCCAAATCTTAAGTATCTC 624  
 DB GCAAGCTCAATGATCTGCTTTCTTAAGAGAAATCTCCAAATCTTAAGTATCTC 624  
 QY 601 AATATTACAGCAGGCTCTAGACAACTTATTAATTTGCTGATGCTTGACCTCAC 660  
 DB AATATTACAGCAGGCTCTAGACAACTTATTAATTTGCTGATGCTTGACCTCAC 660  
 QY 625 AATATTACAGCAGGCTCTAGACAACTTATTAATTTGCTGATGCTTGACCTCAC 660  
 DB AATATTACAGCAGGCTCTAGACAACTTATTAATTTGCTGATGCTTGACCTCAC 660  
 QY 661 CTGATATTTGCTTAAATGCACTGCTGATGCTTAAATCCCAATTAATCTGGAACAGTTCTAGT 720  
 DB CTGATATTTGCTTAAATGCACTGCTGATGCTTAAATCCCAATTAATCTGGAACAGTTCTAGT 720  
 QY 681 GCTGATATTTGCTTAAATGCACTGCTGATGCTTAAATCCCAATTAATCTGGAACAGTTCTAGT 720  
 DB GCTGATATTTGCTTAAATGCACTGCTGATGCTTAAATCCCAATTAATCTGGAACAGTTCTAGT 720  
 QY 721 GCTGATATTTGCTTAAATGCACTGCTGATGCTTAAATCCCAATTAATCTGGAACAGTTCTAGT 720  
 DB GCTGATATTTGCTTAAATGCACTGCTGATGCTTAAATCCCAATTAATCTGGAACAGTTCTAGT 720  
 QY 781 AATGAGCCAAATACTATCGGACCATGATGCTGCGGAGTAAATGCGACAGTTGGGA 840  
 DB AATGAGCCAAATACTATCGGACCATGATGCTGCGGAGTAAATGCGACAGTTGGGA 840  
 QY 841 AAGGATTAATCACTGAGAGAGCTGTTGCAAGCCATCCGGATTTTTCAGAGCCAGC 900  
 DB AAGGATTAATCACTGAGAGAGCTGTTGCAAGCCATCCGGATTTTTCAGAGCCAGC 900  
 QY 901 TTATATGCTCTAATATTTGCGCGCGCGAGAAAGATGCAATCGCCCTCTAAGTGAATTC 960  
 DB TTATATGCTCTAATATTTGCGCGCGCGAGAAAGATGCAATCGCCCTCTAAGTGAATTC 960  
 QY 961 ATGAGGCTGAGAGAGTCACTGATGAGCACTTCTGCGCAACTTGTCTACATTTAGTGC 1020  
 DB ATGAGGCTGAGAGAGTCACTGATGAGCACTTCTGCGCAACTTGTCTACATTTAGTGC 1020  
 QY 811 ATGAGGCTGAGAGAGTCACTGATGAGCACTTCTGCGCAACTTGTCTACATTTAGTGC 870  
 DB ATGAGGCTGAGAGAGTCACTGATGAGCACTTCTGCGCAACTTGTCTACATTTAGTGC 870

QY 1021 CGGCTGCTCAAGTATGATGATCTTCTGAAAACCTCGCTGTAGACACACTCTGACCCAG 1080  
 DB CGGCTGCTCAAGTATGATGATCTTCTGAAAACCTCGCTGTAGACACACTCTGACCCAG 930  
 QY 1081 ATTAGAAAATTTCAGAAAGTGTATATACATACATCTCCAGAAAAGAGATTTGGCTTGA 1140  
 DB ATTAGAAAATTTCAGAAAGTGTATATACATACATCTCCAGAAAAGAGATTTGGCTTGA 990  
 QY 991 ATTAGAAAATTTCAGAAAGTGTATATACATACATCTCCAGAAAAGAGATTTGGCTTGA 990  
 DB ATTAGAAAATTTCAGAAAGTGTATATACATACATCTCCAGAAAAGAGATTTGGCTTGA 990  
 QY 1141 GGTGTGTGACCACTTCACTGAGAGCAAAATATTTTAACTTAACTTAACTTAACTTAA 1200  
 DB GGTGTGTGACCACTTCACTGAGAGCAAAATATTTTAACTTAACTTAACTTAACTTAA 1050  
 QY 1201 TTCTTAATGTTGAACCTTAAAGATGCTGAGGCAATGAGGCAATGAGGCAATGAGGCA 1260  
 DB TTCTTAATGTTGAACCTTAAAGATGCTGAGGCAATGAGGCAATGAGGCAATGAGGCA 1110  
 QY 1051 TTCTTAATGTTGAACCTTAAAGATGCTGAGGCAATGAGGCAATGAGGCAATGAGGCA 1110  
 DB TTCTTAATGTTGAACCTTAAAGATGCTGAGGCAATGAGGCAATGAGGCAATGAGGCA 1110  
 QY 1261 CACTCATTTTGTGACATGATATCAATCACTGCTGAGACAGAAATTTTAACTTAACTTAA 1320  
 DB CACTCATTTTGTGACATGATATCAATCACTGCTGAGACAGAAATTTTAACTTAACTTAA 1170  
 QY 1111 CACTCATTTTGTGACATGATATCAATCACTGCTGAGACAGAAATTTTAACTTAACTTAA 1170  
 DB CACTCATTTTGTGACATGATATCAATCACTGCTGAGACAGAAATTTTAACTTAACTTAA 1170  
 QY 1321 GACTATGCTGCTCTCTCTCTCTTCTTAAAGGCTGATGCGCCCAAAAGTGGCTGTCAT 1380  
 DB GACTATGCTGCTCTCTCTCTCTTCTTAAAGGCTGATGCGCCCAAAAGTGGCTGTCAT 1230  
 QY 1171 GACTATGCTGCTCTCTCTCTCTTCTTAAAGGCTGATGCGCCCAAAAGTGGCTGTCAT 1230  
 DB GACTATGCTGCTCTCTCTCTCTTCTTAAAGGCTGATGCGCCCAAAAGTGGCTGTCAT 1230  
 QY 1381 GTGGCTGCTGCTCAAGGCAAGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440  
 DB GTGGCTGCTGCTCAAGGCAAGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1290  
 QY 1231 GTGGCTGCTGCTCAAGGCAAGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1290  
 DB GTGGCTGCTGCTCAAGGCAAGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1290  
 QY 1441 TATGCTCACTGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1500  
 DB TATGCTCACTGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1350  
 QY 1291 TATGCTCACTGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1350  
 DB TATGCTCACTGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1350  
 QY 1501 ATCATCACTGCAATGATATCAAGAAAGAAATCAAGCTGCTGCTGCTGCTGCTGCTGCT 1560  
 DB ATCATCACTGCAATGATATCAAGAAAGAAATCAAGCTGCTGCTGCTGCTGCTGCTGCT 1410  
 QY 1351 ATCATCACTGCAATGATATCAAGAAAGAAATCAAGCTGCTGCTGCTGCTGCTGCTGCT 1410  
 DB ATCATCACTGCAATGATATCAAGAAAGAAATCAAGCTGCTGCTGCTGCTGCTGCTGCT 1410  
 QY 1561 CTGGTTCAACAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620  
 DB CTGGTTCAACAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1470  
 QY 1411 CTGGTTCAACAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1470  
 DB CTGGTTCAACAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1470  
 QY 1621 CACTGAATGAGCAGCCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1680  
 DB CACTGAATGAGCAGCCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1530  
 QY 1471 CACTGAATGAGCAGCCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1530  
 DB CACTGAATGAGCAGCCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1530  
 QY 1681 CGCCCCCTTGGGCGGCGCGGCAATGCTATCCCTCCAGTCAACATGAGGCTTTTATG 1740  
 DB CGCCCCCTTGGGCGGCGGCGGCAATGCTATCCCTCCAGTCAACATGAGGCTTTTATG 1590  
 QY 1531 CGCCCCCTTGGGCGGCGGCGGCAATGCTATCCCTCCAGTCAACATGAGGCTTTTATG 1590  
 DB CGCCCCCTTGGGCGGCGGCGGCAATGCTATCCCTCCAGTCAACATGAGGCTTTTATG 1590  
 QY 1741 GTCAAGAAATGTCATGCTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1779  
 DB GTCAAGAAATGTCATGCTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1629

## RESULT 5

US-10-177-245A-3

Sequence 3, Application US/10177245A  
 Publication No. US20030083254A1  
 GENERAL INFORMATION:  
 APPLICANT: McKenzie, Edward Alexander  
 APPLICANT: Stamps, Alasdair Craig  
 APPLICANT: Terrett, Jonathan Alexander  
 APPLICANT: Tyson, Kerry Louise  
 TITLE OF INVENTION: Substances  
 FILE REFERENCE: 2543-1-027  
 CURRENT APPLICATION NUMBER: US/10/177,245A  
 PRIOR FILING DATE: 2002-08-29  
 PRIOR APPLICATION NUMBER: PCT/GB00/04963  
 PRIOR FILING DATE: 2000-12-21  
 PRIOR APPLICATION NUMBER: UK 0008713.0  
 PRIOR FILING DATE: 2000-04-07  
 PRIOR APPLICATION NUMBER: UK 9930392.7  
 PRIOR FILING DATE: 1999-12-22

NUMBER OF SEQ ID NOS: 23  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO: 3  
 LENGTH: 2462  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (1)-(2462)  
 OTHER INFORMATION: n = A,T,C or G  
 US-10-177-245A-3

Query Match 79.7%; Score 1417.8; DB 15; Length 2462;  
 Best Local Similarity 90.1%; Pred. No. 0;

Matches 1603; Conservative 0; Mismatches 2; Indels 174; Gaps 1;

1 ATGAGGCTGCTTGTGCTTCCCTTCCCTGAGGCAATGCTCCGAGCACTCCCGCCCGCCGCG 60  
 601 ATGAGGCTGCTTGTGCTTCCCTTCCCTGAGGCAATGCTCCGAGCACTCCCGCCCGCCGCG 660  
 61 TGCCTAGCCCGCGGCGCTCTCTACTTGGCTGTGCTCCATCTCTCCCTTCTCTCCAG 120  
 661 TGCCTAGCCCGCGGCGCTCTCTACTTGGCTGTGCTCCATCTCTCCCTTCTCTCCAG 720  
 121 GCTGAGACAGAGACCTTGGCTGTGAGACAGAGCTGAGGTTTGAAGAAAAGACCTG 180  
 721 GCTGAGACAGAGACCTTGGCTGTGAGACAGAGCTGAGGTTTGAAGAAAAGACCTG 780  
 181 ATTCTACTGATGTGAGACCAAGAACCAAGTCAAGACATGAGACATCTCTCT 240  
 781 ATTCTACTGATGTGAGACCAAGAACCAAGTCAAGACATGAGACATCTCTCT 840  
 241 CTGACAGCTGATCGCTCATCATTCATGATGCTGCTGCTGCTTCTTAAAGCTCCAGGC 300  
 841 CTGACAGCTGATCGCTCATCATTCATGATGCTGCTGCTGCTTCTTAAAGCTCCAGGC 900  
 301 TTGGTACCTCGGCGCGGGAATTTGGCCGCTTCTGCGCTCGGCGGCAAAAGGAC 360  
 901 TTGGTACCTCGGCGCGGGAATTTGGCCGCTTCTGCGCTCGGCGGCAAAAGGAC 960  
 361 GACTTCCTGAGTTCAGAACTGAGAAACCCGCGCAAAAGCTCGCGGCGCGCGCG 420  
 961 GACTTCCTGAGTTCAGAACTGAGAAACCCGCGCAAAAGCTCGCGGCGCGCGCG 1020  
 421 GATTACTATCTCAAAACCTATGAGATGATGTTGCAAGTGTGCTTAAAGTAA 480  
 1021 GATTACTATCTCAAAACCTATGAGATGATGTTGCAAGTGTGCTTAAAGTAA 1080  
 481 CAGAAAGCTGCAAGATGCTCCAGCACTGATGTTATGCTGCTCCAAAGGAGAG 540  
 1081 CAGAAAGCTGCAAGATGCTCCAGCACTGATGTTATGCTGCTCCAAAGGAGAG 1140  
 541 GCGAGCTAGATGATCTGTTCTTTTAAAGAGACATTTCCCAATCTTAAGTATCTC 600  
 1141 GCGAGCTAGATGATCTGTTCTTTTAAAGAGACATTTCCCAATCTTAAGTATCTC 1200  
 601 ATATTAACAGCAGCTCTAGACAAACTTAACTTATGCTGATGCTCTGAGCTCAC 660  
 1201 ATATTAACAGCAGCTCTAGACAAACTTAACTTATGCTGATGCTCTGAGCTCAC 1209  
 661 CTGATATTTGCTCTAAATGACATGCGCTGTAATCCCAATTACTCTGAGACAGTTCTAGT 720  
 1210 ----- 1209  
 721 GCGCTAGTCTGTGAAGTACAGCGCAGCAAAAGTAAACATTTCTGGAGACTGGGT 780  
 1210 ----- 1209  
 781 AATGAGCAAAATTAATCTATGAGACATGATGCGCGGAGTAAATGAGACAGCTTGGGA 840  
 1210 -----GAGCAAAATTAATCTATGAGACATGATGCGCGGAGTAAATGAGACAGCTTGGGA 1266  
 841 AAGGATTAATCAATGAGTGAAGAGCTGTGACAGCCCATCGGATTTATTCAGAGCCAGC 900

1267 AAGGATTAATCAATGAGTGAAGAGCTGTGACAGCCCATCGGATTTATTCAGAGCCAGC 1326  
 901 TTAATAGCCCTTAATTTGGCGCGCGGAGAAAGATGATGCGCCCTCTAGATGATTC 960  
 1327 TTAATAGCCCTTAATTTGGCGCGCGGAGAAAGATGATGCGCCCTCTAGATGATTC 1386  
 961 ATGAGGTGACAGAAATGACATGATGAGTGAAGTTACTCTGCAACATTTGCTACATGAGTC 1020  
 1387 ATGAGGTGACAGAAATGACATGATGAGTGAAGTTACTCTGCAACATTTGCTACATGAGTC 1446  
 1021 CGGCTGCTCAAGTGAATGACATTTCTGAAAACCTCGCTTTTGAACAATCTCTGACAG 1080  
 1447 CGGCTGCTCAAGTGAATGACATTTCTGAAAACCTCGCTTTTGAACAATCTCTGACAG 1506  
 1081 ATTAGCAAAATTCAGAAAGTGTAAATATACATACATCTCGAGAAAGAAATTTGGCTTGA 1140  
 1507 ATTAGCAAAATTCAGAAAGTGTAAATATACATACATCTCGAGAAAGAAATTTGGCTTGA 1566  
 1141 GGTGTGTGACACCTGAGCTGAGGACAAACATCTATCCGATTTCTATGCTGACAGA 1200  
 1567 GGTGTGTGACACCTGAGCTGAGGACAAACATCTATCCGATTTCTATGCTGACAGA 1626  
 1201 TTCTTATGTTGAACCTTAGGATGCTGAGCAATCAAGGCAATGATGCTGATACCG 1260  
 1627 TTCTTATGTTGAACCTTAGGATGCTGAGCAATCAAGGCAATGATGCTGATACCG 1686  
 1261 CACTCATTTTGAACATGAGATCAATCACTCTGTGAGCAACGAATTTTAAACCATTTACA 1320  
 1687 CACTCATTTTGAACATGAGATCAATCACTCTGTGAGCAACGAATTTTAAACCATTTACA 1746  
 1321 GACTACGCTCTCTCTCTCTCAAGCGCTGATAGGCGCCCAAGTCTGCTGATGAT 1380  
 1747 GACTACGCTCTCTCTCTCTCAAGCGCTGATAGGCGCCCAAGTCTGCTGATGAT 1806  
 1381 GTGCTGAGGCTGACAGCGGAAGCCAGCGCTGAGGATGACCGGACAAACTTAAGATT 1440  
 1807 GTGCTGAGGCTGACAGCGGAAGCCAGCGCTGAGGATGACCGGACAAACTTAAGATT 1866  
 1441 TATGCTACTGACAAACCAACCAACCAACCAATAGCTGTGAGGCTCACTTACCTTTT 1500  
 1867 TATGCTACTGACAAACCAACCAACCAACCAACCAATAGCTGTGAGGCTCACTTACCTTTT 1926  
 1501 ATCATCACTTGTCACTCATCAAGAAAGAAATCAAGCTGCTGAGACTCTCAGAGCAAG 1560  
 1927 ATCATCACTTGTCACTCATCAAGAAAGAAATCAAGCTGCTGAGACTCTCAGAGCAAG 1986  
 1561 CTGCTTACCAATGCTGCTGACAGCTTATGAGGCAAGAGGCTTAAAGTCAAGTCAAGT 1620  
 1987 CTGCTTACCAATGCTGCTGACAGCTTATGAGGCAAGAGGCTTAAAGTCAAGTCAAGT 2046  
 1621 CAATGATAGGCGAGCCCTTAGTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680  
 2047 CAATGATAGGCGAGCCCTTAGTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2106  
 1681 CGCCCTCTGCGGCGCGGCGGCAATGATGATCTCTCAATCACTCAATGAGCTTTTATG 1740  
 2107 CGCCCTCTGCGGCGCGGCGGCAATGATGATCTCTCAATCACTCAATGAGCTTTTATG 2166  
 1741 GTCAAGATGATGATGCTTGGCTGCGGCTAGCGATTA 1779  
 2167 GTCAAGATGATGATGCTTGGCTGCGGCTAGCGATTA 2205

# RESULT 6

US-09-880-262-1

; Sequence 1, Application US/09880262

; Patent No. US20020137907A1

; GENERAL INFORMATION:

; APPLICANT: SOUTHAN, CHRISTOPHER DONALD

; APPLICANT: LARMINIE, CHRISTOPHER GEORFREY CARSON

; APPLICANT: RANCE, KIM

; APPLICANT: HAYES, PHILIP DAVID

TITLE OF INVENTION: NOVEL COMPOUNDS  
 FILE REFERENCE: GP-30231  
 CURRENT APPLICATION NUMBER: US/09/880,262  
 CURRENT FILING DATE: 2001-06-13  
 PRIOR APPLICATION NUMBER: UK 0014447.7  
 PRIOR FILING DATE: 2000-06-13  
 NUMBER OF SEQ ID NOS: 2  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 1  
 LENGTH: 1602  
 TYPE: DNA  
 ORGANISM: HOMO SAPIENS  
 US-09-880-262-1

Query Match 79.6%; Score 1416.4; DB 9; Length 1602;  
 Best Local Similarity 90.1%; Pred. No. 0;  
 Matches 1601; Conservative 0; Mismatches 1; Indels 174; Gaps 1;

QY 1 ATGAGGGTGTCTGTGCTTCCCTGAAAGCCATGCGCCCTCCAGCAACTCCGCCCCCGCG 60  
 DB 1 ATGAGGGTGTCTGTGCTTCCCTGAAAGCCATGCGCCCTCCAGCAACTCCGCCCCCGCG 60  
 QY 61 TGCCTAGCCCGGCGGCTCTCTAATTGCTCTGCTGCTCTCTCTCTCTCTCTCTCTCT 120  
 DB 61 TGCCTAGCCCGGCGGCTCTCTAATTGCTCTGCTGCTCTCTCTCTCTCTCTCTCTCT 120  
 QY 121 GCTGAGACAGAGAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
 DB 121 GCTGAGACAGAGAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
 QY 181 ATTCTACTGATGAGACCAAGAACCCAGTCAGACAGTCATGAGAACTTCTCTCT 240  
 DB 181 ATTCTACTGATGAGACCAAGAACCCAGTCAGACAGTCATGAGAACTTCTCTCT 240  
 QY 241 CTGCAAGTGAATCCCTGCAATTCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
 DB 241 CTGCAAGTGAATCCCTGCAATTCATGATGCTGCTGCTGCTGCTGCTGCTGCTG 300  
 QY 301 TTGGTGAACCTGCGCGGCGGACTTTTGGCGGCTTTCTGCGCTTCCGCGGCAAAAGAC 360  
 DB 301 TTGGTGAACCTGCGCGGCGGACTTTTGGCGGCTTTCTGCGCTTCCGCGGCAAAAGAC 360  
 QY 361 GACTTCTGCAATTCAGAACCTGAGAACCCGCGGCGGCGGCGGCGGCGGCGGCGG 420  
 DB 361 GACTTCTGCAATTCAGAACCTGAGAACCCGCGGCGGCGGCGGCGGCGGCGGCGG 420  
 QY 421 GATTACTATCTCAAAAATCTAGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
 DB 421 GATTACTATCTCAAAAATCTAGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
 QY 481 CAGAAAGCTGCAAGATGCTCCAGCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
 DB 481 CAGAAAGCTGCAAGATGCTCCAGCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
 QY 541 GCAAGCTGCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
 DB 541 GCAAGCTGCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
 QY 601 AATATTAAGCCAGGCTCTAGCAAACTTTATTAATTGCTGCTGCTGCTGCTGCTGCT 660  
 DB 601 AATATTAAGCCAGGCTCTAGCAAACTTTATTAATTGCTGCTGCTGCTGCTGCTGCT 660  
 QY 661 CTGATATTGCTCTAATGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
 DB 661 CTGATATTGCTCTAATGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
 QY 721 GCCCTGAGTGTGTAAGTACAGCGCCAGCAAAAGTAACTTTCTTGGAACTGGGT 780  
 DB 721 GCCCTGAGTGTGTAAGTACAGCGCCAGCAAAAGTAACTTTCTTGGAACTGGGT 780  
 QY 781 AATGAGCCAAATACTATGCAATGCAATGCGCGGCAATTAATGCAAGCACTGTTGGA 840  
 DB 781 AATGAGCCAAATACTATGCAATGCAATGCGCGGCAATTAATGCAAGCACTGTTGGA 840  
 QY 840 ---GAGCAATACTATGCAATGCAATGCGCGGCAATTAATGCAAGCACTGTTGGA 866  
 DB 840 ---GAGCAATACTATGCAATGCAATGCGCGGCAATTAATGCAAGCACTGTTGGA 866

QY 841 AAGATTATCATCCAGCTGGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
 DB 841 AAGATTATCATCCAGCTGGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
 QY 901 TTATATGAGCCCTAATATTTGGGCGGCGGAGGAGAAATGCTAGCGCTCTAGATGATTC 960  
 DB 901 TTATATGAGCCCTAATATTTGGGCGGCGGAGGAGAAATGCTAGCGCTCTAGATGATTC 960  
 QY 961 ATGAGAGTGCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
 DB 961 ATGAGAGTGCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
 QY 1021 CGGGTGTGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080  
 DB 1021 CGGGTGTGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080  
 QY 1081 ATTAGAAATTTCAAGAAAGTGTATATATATATATATATATATATATATATATAT 1140  
 DB 1081 ATTAGAAATTTCAAGAAAGTGTATATATATATATATATATATATATATATATAT 1140  
 QY 1141 GGTGTGTGACACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
 DB 1141 GGTGTGTGACACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
 QY 1201 TTCTTATGTTGAACATTTTGAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
 DB 1201 TTCTTATGTTGAACATTTTGAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
 QY 1261 CACTCATTTTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320  
 DB 1261 CACTCATTTTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320  
 QY 1321 GACTACTGCT 1380  
 DB 1321 GACTACTGCT 1380  
 QY 1381 GTGGCTGGGCTCCAGCGGAGCCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440  
 DB 1381 GTGGCTGGGCTCCAGCGGAGCCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440  
 QY 1441 TATGCTACCTGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1500  
 DB 1441 TATGCTACCTGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1500  
 QY 1501 ATCATCACTTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560  
 DB 1501 ATCATCACTTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560  
 QY 1561 CTGGTTCACAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620  
 DB 1561 CTGGTTCACAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620  
 QY 1621 CAAGTGAATGCAAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATG 1680  
 DB 1621 CAAGTGAATGCAAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATG 1680  
 QY 1681 CGCCCTTGGGCGGCGGCGGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740  
 DB 1681 CGCCCTTGGGCGGCGGCGGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740  
 QY 1741 GTCAAGATGCTCAATGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800  
 DB 1741 GTCAAGATGCTCAATGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800  
 QY 1801 GTCAAGATGCTCAATGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860  
 DB 1801 GTCAAGATGCTCAATGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860

RESULT 7  
 US-09-978-249-2  
 ; Sequence 2, Application US/09978249  
 ; Patent No. US20020106780A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pisciella, et al.  
 ; TITLE OF INVENTION: Extracellular Matrix Polynucleotides, Polypeptides, and Antibodies

FILE REFERENCE: PT054P1  
 CURRENT APPLICATION NUMBER: US/09/978,249  
 CURRENT FILING DATE: 2001-10-17  
 PRIOR APPLICATION NUMBER: PCT/US01/11643  
 PRIOR FILING DATE: 2001-04-11  
 PRIOR APPLICATION NUMBER: 60/198,123  
 PRIOR FILING DATE: 2000-04-18  
 NUMBER OF SEQ ID NOS: 16  
 SOFTWARE: Patent Ver. 2.0  
 SEQ ID NO: 2  
 LENGTH: 1930  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-09-978-249-2

Query Match 61.6%; Score 1095.4; DB 9; Length 1930;  
 Best Local Similarity 81.1%; Pred. No. 0;  
 Matches 1442; Conservative 0; Mismatches 1; Indels 336; Gaps 1;

QY 1 ATGAGGCTGCTTGTGCTTCCCTGAAGCCATGCGCCGCAACTCCGCCCCCGCG 60  
 DB 19 ATGAGGCTGCTTGTGCTTCCCTGAAGCCATGCGCCGCAACTCCGCCCCCGCG 78  
 QY 61 TGCCTGAGCCCGGCGGCTCTCTAATTGCTTGTGCTCACTCTCCCTTCTCCAG 120  
 DB 79 TGCCTGAGCCCGGCGGCTCTCTAATTGCTTGTGCTCACTCTCCCTTCTCCAG 138  
 QY 121 GCTGAGACAGAGAGACCTTGCCTGTAGACAGAGCTGAGGTTGAGAGAAAGACCTG 180  
 DB 139 GCTGAGACAGAGAGACCTTGCCTGTAGACAGAGCTGAGGTTGAGAGAAAGACCTG 198  
 QY 181 ATTCTATGATGTAGACACCAAGAACCCAGTCAGACACTCAATAGAACTTCTCT 240  
 DB 199 ATTCTATGATGTAGACACCAAGAACCCAGTCAGACACTCAATAGAACTTCTCT 258  
 QY 241 CTGAGCTGATCCGTCCATCATTCATGATGCTGCTGATTTCTTAAGCTCCAGCG 300  
 DB 259 CTGAGCTGATCCGTCCATCATTCATGATGCTGCTGATTTCTTAAGCTCCAGCG 318  
 QY 301 TTGCTGACCTTGGCCCGGAGCTTGGCCGCTTTCTGCTTCCGCGGCAAAAGACC 360  
 DB 319 TTGCTGACCTTGGCCCGGAGCTTGGCCGCTTTCTGCTTCCGCGGCAAAAGACC 378  
 QY 361 GACTCTCGATGCTCCGAGACCTGAGAAACCGCGGAGAAACCGCGGCGCGCG 420  
 DB 379 GACTCTCGATGCTCCGAGACCTGAGAAACCGCGGAGAAACCGCGGCGCGCG 438  
 QY 421 GATTACTATCTCAAAAATCTAGAGATGACATTTGCAAGTATGTTGCTTATGATAA 480  
 DB 439 GATTACTATCTCAAAAATCTAGAG----- 463  
 QY 481 CAGAAAGCTGCAAGATGGCCGAGACCTGATGTTATGCTGCTCCAAAGGAGAG 540  
 DB 464 ----- 463  
 QY 541 GCAGCTCAGATGATCTGTTCTTCTTAAGAGCAATTCCTCAATCTTACGATATCT 600  
 DB 464 ----- 463  
 QY 601 ATTCTAAGCAGGCTCTAGACAAACCTTATTAATTCTGATGCTCTGAGCTCCAC 660  
 DB 464 ----- 463  
 QY 661 CTGATATTTGCTTAATGACATGCTGATATCCCAATACTCTGAAACAGTTCTAG 720  
 DB 464 ----- 463  
 QY 721 GCCCTGAGCTGTGAAAGTACAGGCGCAGAAAAAGTACAACATTTCTGGGAACGCG 780  
 DB 464 ----- 463  
 QY 781 AATGAGCAATACTATCGACCATGATGCGCGGCACTTAATGAGCAGCTTGGGAG 840  
 DB ----- 463

DB 464 -ATGAGCAATACTATCGACCATGATGCGCGGCACTTAATGAGCAGCTTGGGAG 522  
 QY 841 AAGATTACATCCAGCTGAAGAGCTTGTGCAAGCCATCCGATTTATTCAGAGCCAG 900  
 DB 523 AAGATTACATCCAGCTGAAGAGCTTGTGCAAGCCATCCGATTTATTCAGAGCCAG 582  
 QY 901 TTATATGCCCCATATTTGGGCGGCGAGAGAAATGCTATGCGCCCTAGATGATTC 960  
 DB 583 TTATATGCCCCATATTTGGGCGGCGAGAGAAATGCTATGCGCCCTAGATGATTC 642  
 QY 961 ATGAAAGTGCAGAGAGTACAGTATGATGATTAACCTGCGCAACTTGCATGATGCG 1020  
 DB 643 ATGAAAGTGCAGAGAGTACAGTATGATGATTAACCTGCGCAACTTGCATGATGCG 702  
 QY 1021 CCGGTGTCAGAGTATGATGATTAACCTGCGCAACTTGCATGATGATGCG 1080  
 DB 703 CCGGTGTCAGAGTATGATGATTAACCTGCGCAACTTGCATGATGATGCG 762  
 QY 1081 ATTAGAAAATTCAGAAAAGTGTATATATACATCTCCAGAAAAGAGATTGGCTTGA 1140  
 DB 763 ATTAGAAAATTCAGAAAAGTGTATATATACATCTCCAGAAAAGAGATTGGCTTGA 822  
 QY 1141 GGTGTGTCAGACCTCAGCTGAGAGGCAAAACATTTATCCGATTTCTATGCTGAGGA 1200  
 DB 823 GGTGTGTCAGACCTCAGCTGAGAGGCAAAACATTTATCCGATTTCTATGCTGAGGA 882  
 QY 1201 TTCTATGTTGAACAATTAGAAATGCTGCGCAATCAGGGCAATGATGCTGATGCG 1260  
 DB 883 TTCTATGTTGAACAATTAGAAATGCTGCGCAATCAGGGCAATGATGCTGATGCG 942  
 QY 1261 CACTCATTTTGTGACATGATATCATCTGCTGTCAGACAGAAATTTAAACCAATTA 1320  
 DB 943 CACTCATTTTGTGACATGATATCATCTGCTGTCAGACAGAAATTTAAACCAATTA 1002  
 QY 1321 GACTACTGCTCTCTCTCTCTCAAGAGCTGATGCGCCCAAGCTTGGCTGTCAT 1380  
 DB 1003 GACTACTGCTCTCTCTCTCTCAAGAGCTGATGCGCCCAAGCTTGGCTGTCAT 1062  
 QY 1381 GTGCTGAGCTCCAGCGAAGCAAGCGCTGCGGAGTATCCGAGCAAACTAAGATT 1440  
 DB 1063 GTGCTGAGCTCCAGCGAAGCAAGCGCTGCGGAGTATCCGAGCAAACTAAGATT 1122  
 QY 1441 TATGCTCAGTCAAAACCAACACACACACACTAGTGTGTCATTTAATCACTTTT 1500  
 DB 1123 TATGCTCAGTCAAAACCAACACACACACACTAGTGTGTCATTTAATCACTTTT 1182  
 QY 1501 ATCATCACTTGCATGATCAAGAAAGAAATCAAGCTGCTGAGACTCTCAGAGCAAG 1560  
 DB 1183 ATCATCACTTGCATGATCAAGAAAGAAATCAAGCTGCTGAGACTCTCAGAGCAAG 1242  
 QY 1561 CTGCTTCAAGTACCTGCTGAGAGCTTATGAGAGAGAGGCTTAAGTCCAGATCAGT 1620  
 DB 1243 CTGCTTCAAGTACCTGCTGAGAGCTTATGAGAGAGAGGCTTAAGTCCAGATCAGT 1302  
 QY 1621 CAAGTATGAGCAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
 DB 1303 CAAGTATGAGCAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1362  
 QY 1681 CGCCCTCTTGGGCGGCGGAGCAATGCTCTCAAGTCAACATGAGGCTTTATGCT 1740  
 DB 1363 CGCCCTCTTGGGCGGCGGAGCAATGCTCTCAAGTCAACATGAGGCTTTATGCT 1422  
 QY 1741 GTCAAGATGTCATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1779  
 DB 1423 GTCAAGATGTCATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1461

RESULT 8  
 US-10-177-245A-5  
 ; Sequence 5, Application US/10177245A  
 ; Publication No. US2003083254A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: McKenzie, Edward Alexander



APPLICANT: Stamps, Alasdair Craig  
 APPLICANT: Terrett, Jonathan Alexander  
 APPLICANT: Tyson, Kerry Louise  
 TITLE OF INVENTION: Substances  
 FILE REFERENCE: 2543-1-027  
 CURRENT APPLICATION NUMBER: US/10/177,245A  
 CURRENT FILING DATE: 2002-08-29  
 PRIOR APPLICATION NUMBER: PCT/GB00/04963  
 PRIOR FILING DATE: 2000-12-21  
 PRIOR APPLICATION NUMBER: UK 0008713.0  
 PRIOR FILING DATE: 2000-04-07  
 PRIOR APPLICATION NUMBER: UK 9930392.7  
 PRIOR FILING DATE: 1999-12-22  
 NUMBER OF SEQ ID NOS: 23  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 5  
 LENGTH: 2300  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: (1)..(2300)  
 OTHER INFORMATION: n = A,T,C or G  
 US-10-177-245A-5

Query March 61.6%; Score 1095.4; DB 15; Length 2300;

Best Local Similarity 81.1%; Pred. No. 0;

Matches 144; Conservative 0; Mismatches 1; Indels 336; Gaps 1;

QY 1 ATGAGGAGTGTGCTTGTGCTTCCCTGAGCCATGCGCCCTCCAGCACTCCGCCCCCGCG 60  
 DB 601 ATGAGGAGTGTGCTTGTGCTTCCCTGAGCCATGCGCCCTCCAGCACTCCGCCCCCGCG 660  
 QY 61 TGCCTACCCCGGGGGCTCTCTACTTGGCTGTGTGCTCATCTTCCCTTTCTCCAG 120  
 DB 661 TGCCTACCCCGGGGGCTCTCTACTTGGCTGTGTGCTCATCTTCCCTTTCTCCAG 720  
 QY 121 GCTGAGACAGAGAGACCTTGGCTGTGAGACAGCTGAGCTGAGTGAAGAAAGACCTG 180  
 DB 721 GCTGAGACAGAGAGACCTTGGCTGTGAGACAGCTGAGCTGAGTGAAGAAAGACCTG 780  
 QY 181 ATTCTACTGATGTGAGACCAAGAACCCAGTCAAGCACTGATGAACTTCTCTCT 240  
 DB 781 ATTCTACTGATGTGAGACCAAGAACCCAGTCAAGCACTGATGAACTTCTCTCT 840  
 QY 241 CTGAGCTGATCCGCTCATCATTCATGATGAGCTGAGTGTGATTCCTAAGCTCCAGCG 300  
 DB 841 CTGAGCTGATCCGCTCATCATTCATGATGAGCTGAGTGTGATTCCTAAGCTCCAGCG 900  
 QY 301 TTGGTGAACCTTGGCCCGGAGACTTTCGCGCCCTTTCGTGCTTGGGGGCAAAAGACG 360  
 DB 901 TTGGTGAACCTTGGCCCGGAGACTTTCGCGCCCTTTCGTGCTTGGGGGCAAAAGACG 960  
 QY 361 GACTCTCTGAGTTCGAGAACCTGAGAACCCGGCGGAAAGCCGGGGGCGCGGCGCG 420  
 DB 961 GACTCTCTGAGTTCGAGAACCTGAGAACCCGGCGGAAAGCCGGGGGCGCGGCGCG 1020  
 QY 421 GATTACTATCTCAAAAATCTAGAGATGACATTTGGAAGTATGTTGCTTGAATAA 480  
 DB 1021 GATTACTATCTCAAAAATCTAGAGATGACATTTGGAAGTATGTTGCTTGAATAA 1045  
 QY 481 CAGAAAGCTGCAAGATTTGCCAGACACCTGATGTTATGCTGTGCTCCAAAGGAGAG 540  
 DB 1046 ----- 1045  
 QY 541 GCAAGCTCAGATGACATCTGTTCTTCTAAGAGACATTTCTCAATCTTACAGTAATCTC 600  
 DB 1046 ----- 1045  
 QY 601 ATATTACAGCCAGGCTCTTACAGAACTTATACTTGTGATGTTGCTGTGACTCCAC 660  
 DB 1046 ----- 1045

QY 661 CTGATATTGCTTAAATGACATGCGCTGATATCCCAATTAATCTCTGGAACTGTTACT 720  
 DB 1046 ----- 1045  
 QY 721 GCCCTGAGTCTGTGAAGTACAGGCCAGCAAAAAGTACAAATTTCTTGGGAATGGGT 780  
 DB 1046 ----- 1045  
 QY 781 AATGAGCCAAATTAATCTAGAGCAATGAGCCAGTGAATGAGCAAGCTTGGAG 840  
 DB 1046 -ATGAGCCAAATTAATCTAGAGCAATGAGCCAGTGAATGAGCAAGCTTGGAG 1104  
 QY 841 AAGATTTACATCCAGCTGAAAGAGCTTGTGAGCCATCCGGAATTTTTCAGAGCCAGC 900  
 DB 1105 AAGATTTACATCCAGCTGAAAGAGCTTGTGAGCCATCCGGAATTTTTCAGAGCCAGC 1164  
 QY 901 TTAATGAGCCCTTAATATTGAGCGGCGCGAGAGAAATGTCAATGCGCTCTAGATGATTC 960  
 DB 1165 TTAATGAGCCCTTAATATTGAGCGGCGCGAGAGAAATGTCAATGCGCTCTAGATGATTC 1224  
 QY 961 ATGAAAGTGGCAGAGATGACATGATGCACTTCTGGCAACCTTGTCAATTTAGTGC 1020  
 DB 1225 ATGAAAGTGGCAGAGATGACATGATGCACTTCTGGCAACCTTGTCAATTTAGTGC 1284  
 QY 1021 CGGAGTGTCAAGTGTAGATGATCTTCTGAAACTGCGCTGTGAGACACATCTCTGACAG 1080  
 DB 1285 CGGAGTGTCAAGTGTAGATGATCTTCTGAAACTGCGCTGTGAGACACATCTCTGACAG 1344  
 QY 1081 ATTAGAAATTTCAAGAAATGTTTATATATCACTCACTCAAGAAAGAAATTTGGCTTGA 1140  
 DB 1345 ATTAGAAATTTCAAGAAATGTTTATATATCACTCACTCAAGAAAGAAATTTGGCTTGA 1404  
 QY 1141 GGTGTGTGACCACTCACTGAGAGGAGCAAAACATCTATGATGATCTGATGATGATG 1200  
 DB 1405 GGTGTGTGACCACTCACTGAGAGGAGCAAAACATCTATGATGATGATGATGATGATG 1464  
 QY 1201 TTTCTATGTTGAACATTTTGAAGATGTGGCAATCAAGGCAATGATGTGTGATACGG 1260  
 DB 1465 TTTCTATGTTGAACATTTTGAAGATGTGGCAATCAAGGCAATGATGTGTGATACGG 1524  
 QY 1261 CACTCATTTTGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320  
 DB 1525 CACTCATTTTGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1584  
 QY 1321 GACTACTGCT 1380  
 DB 1585 GACTACTGCT 1644  
 QY 1381 GTGGCTGGGCTTCAAGGAGACCAAGGCTTGGCGAGATGATGCGGAGCAAACTAAGAT 1440  
 DB 1645 GTGGCTGGGCTTCAAGGAGACCAAGGCTTGGCGAGATGATGCGGAGCAAACTAAGAT 1704  
 QY 1441 TATGCTCACTGACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1500  
 DB 1705 TATGCTCACTGACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1764  
 QY 1501 ATCATCACTTGCATGATCAAGAAAGAAATCAAGCTGCTGGAGCTCTGAGACAG 1560  
 DB 1765 ATCATCACTTGCATGATCAAGAAAGAAATCAAGCTGCTGGAGCTCTGAGACAG 1824  
 QY 1561 CTGTTCAACAGTATCTGCTGAGCGCTATGAGGAGAGAGGAGGAGGAGGAGGAGGAG 1620  
 DB 1825 CTGTTCAACAGTATCTGCTGAGCGCTATGAGGAGAGAGGAGGAGGAGGAGGAGGAG 1884  
 QY 1621 CAATGAATGAGCAAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATG 1680  
 DB 1885 CAATGAATGAGCAAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATG 1944  
 QY 1681 GGGCGGCTTGGGCGGCGGAGCAATGATGATGATGATGATGATGATGATGATGATGATG 1740  
 DB 1945 GGGCGGCTTGGGCGGCGGAGCAATGATGATGATGATGATGATGATGATGATGATGATG 2004  
 QY 1741 GTCAAGAAATGTCATGCTTGGCTGCGCTACGATTA 1779

Db 2005 GTCAAGATGTCATGCTTGGCCTGCGGCTACCGATTA 2043

```

RESULT 9
US-10-080-254-53
: Sequence 53, Application US/10080254
: Publication No. US20030199008A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: P1114C1
: CURRENT APPLICATION NUMBER: US/10/080,254
: CURRENT FILING DATE: 2002-02-22
: NUMBER OF SEQ ID NOS: 169
: Prior Application removed - See File Wrapper or Palm
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 53
: LENGTH: 1114
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-080-254-53

```

Query Match	35.5%;	Score 632.4;	DB 15;	Length 1114;
Best Local Similarity	99.8%;	Pred. No. 5.8e-199;		
Matches 633;	Conservative,	0;	Mismatches 1;	Indels 0;
				Gaps 0;

QY	1146	AGTAGCAACCTCAGCTGAGGACAAACAAATCTATCCCATTTCTATGCTGACAGATTCTT	1205
Db	12	GGTAGCAACCTCAGCTGAGGACAAACAAATCTATCCCATTTCTATGCTGACAGATTCTT	71
QY	1206	ATGTGTAAACACTTTAGGATATCTGTCGCAATACAGGGCAATTGATGTGTGATACGGCACTC	1265
Db	72	ATGTGTAAACACTTTAGGATATGCTGTGCCAAATACAGGGCAATTGATGTGTGATACGGCACTC	131
QY	1266	ATTTTATGACATGATATCAATCATCCTGTGACACAGATTTTAACCATTAACGACATA	1325
Db	132	ATTTTATGACATGATATCAATCATCCTGTGACACAGATTTTAACCATTAACGACATA	191
QY	1326	CTGGCTCTCTCTCTCTCTAACAAGCCTTATCGGCCCCAAAGCTTTGCTGTGATGTGGC	1385
Db	192	CTGGCTCTCTCTCTCTCTAACAAGCCTTATCGGCCCCAAAGCTTTGCTGTGATGTGGC	251
QY	1386	TGGGCTCCAGCGGAAGCCACGGGCTGGCGAGTATCCGGGCAAACTTAGATTTATG	1445
Db	252	TGGGCTCCAGCGGAAGCCACGGGCTGGCGAGTATCCGGGCAAACTTAGATTTATG	311
QY	1446	TCACTGCAAAACCCACCAACAACAACTAAGTTGTGGGTTCATTAACATTTTATCAT	1505
Db	312	TCACTGCAAAACCCACCAACAACAACAACTAAGTTGTGGGTTCATTAACATTTTATCAT	371
QY	1506	CAACTGTGATGATTAAGAAAGAAATCAAGTGTGCTGGACCTTCACAGACAAGCTGT	1565
Db	372	CAACTGTGATGATTAAGAAAGAAATCAAGTGTGCTGGACCTTCACAGACAAGCTGT	431
QY	1566	TCACCAGTACTCTGTGCAAGCCCTATGGCAAGAGGCTTAAAGTCCAAAGTCAGTCAACT	1625
Db	432	TCACCAGTACTCTGTGCAAGCCCTATGGCAAGAGGCTTAAAGTCCAAAGTCAAGTCAACT	491
QY	1626	GAATGGCCAGGCTTATGATGTGTGACGACGGGACCTTCCACAGATTGAAGCCCGGCGC	1685
Db	492	GAATGGCCAGGCTTATGATGTGTGACGACGGGACCTTCCACAGATTGAAGCCCGGCGC	551
QY	1686	CCTTGGGCGCGGCGGACATTTGTATCCTTCCAGTACCAATGGGCTTTTATGTGTCA	1745
Db	552	CCTTGGGCGCGGCGGACATTTGTATCCTTCCAGTACCAATGGGCTTTTATGTGTCA	611
QY	1746	GAATGTCAATCTTTGGGCTGCGCGTACCGATTA	1779
Db	612	GAATGTCAATCTTTGGGCTGCGCGTACCGATTA	645

US-10-242-355-322  
; Sequence 322, Application US/10242355  
; Publication No. US20030235831A1  
GENERAL INFORMATION:

Query Match	35.5%	Score 632.4	DB 16	Length 114
Best Local Similarity	99.8%	Pred. No. 5.8e-199		
Matches 633	Conservative 0	Mismatches 1	Indels 0	Gaps 0

QY	1146	GGTACACACCGCTCACTGAGAGCAAAACAATCTATCCGATCTCCATGCTGACGAACTTCTT	1205
Dd	12	GGTACACACCTCACTGAGAGCAAAACAATCTATCCGATCTCCATGCTGAGAACTTCTT	71
QY	1206	ATGGTTGAACACTTTAGGAATGCTGGCCAATCAGGGCATTGATGTCGTGATACGGCACTC	1265
Dd	72	ATGGTTGAACCTTTAGGAATGCTGGCCAATCAGGGCATTGATGTCGTGATACGGCACTC	131
QY	1286	ATTTTTTGACCATGGATATACATCACTCTCGTGACACAATTTTAAACCATTTAACACAGACTA	1325
Dd	132	ATTTTTTGACCATGGATATACATCACTCTCGTGACACAATTTTAAACCATTTAACACAGACTA	191
QY	1326	CTGGCTCTCTCTCTCTCTACAAAGCGCTGATCGGCCCAAAAGTCTTGCTGTCATGTCGAC	1385
Dd	192	CTGGCTCTCTCTCTCTCTACAAAGCGCTGATCGGCCCAAAAGTCTTGCTGTCATGTCGAC	251
QY	1386	TGGGCTCGAGCGGAGCGACGGCTTGGCCGAGTGATCCGGGACAACTAAGGATTTTATGC	1445
Dd	252	TGGGCTCGAGCGGAGCGACGGCTTGGCCGAGTGATCCGGGACAACTAAGGATTTTATGC	311
QY	1446	TCACGTGACAAACACACACAACAACAATCACTGTTGAGGTCATTAACCTTTTATCAT	1505
Dd	312	TCACGTGACAAACACACAACAACAATCACTGTTGAGGTCATTAACCTTTTATCAT	371
QY	1506	CAACTTGCATGCATCAAGAAAGAAATCAAGCTTGCTGGACCTCTCAGAGACAAGCTGCT	1565
Dd	372	CAACTTGCATGCATCAAGAAAGAAATCAAGCTTGCTGGACCTCTCAGAGACAAGCTGCT	431
QY	1566	TCACCAAGTACCTGCTGACGCGCTTATGGGACAGAGGGCTTAAAGTCAAGTCAAGGCACT	1625
Dd	432	TCACCAAGTACCTGCTGACGCGCTTATGGGACAGAGGGCTTAAAGTCAAGTCAAGGCACT	491





SOFTWARE: Patentin version 3.1  
 SEQ ID NO 10  
 LENGTH: 1605  
 TYPE: DNA  
 ORGANISM: Gallus gallus  
 US-09-930-218-10

Query Match 16.5%; Score 292.8; DB 9; Length 1605;  
 Best Local Similarity 53.7%; Pred. No. 9.7e-86;  
 Matches 649; Conservative 0; Mismatches 532; Indels 27; Gaps 1;

```

QY 560 TTCTTTAAAGGAGCAATTCCTCAATCTTACGATATCTCATATTAAGCAGAGTCTC 619
DB 408 TGTCTCTCGCTGAACATCTCTGAAAAAGACAAAACACCAATTAACAAGAGCAGC 467
QY 620 TAGACAACTTTAACTTTGCTGATTTGCTGACCTCACTGATATTTGCTTAATG 679
DB 468 TGAACATCTCCACAGGTTGCGACAGAGCTCAGGCTTCGGCTGGATTTGGGCTGAAG 527
QY 680 CACTGCTGCTAATCCCAATTAATCTCTGAAACGTTTATGTCCTGAGTCTGTAAGT 739
DB 528 CACTGCTGCTGACAGGCTGAGCTGAGTGGAGACGCTTCAACGCAAGAGCTGAGGCT 587
QY 740 ACAGCGCCAGCAAAAGTATCAACATTTCTTGAGAACTGGATATGAGCCAAATTAATC 799
DB 588 ACTGTGACAGCGCAGCTCAACATCTCTGAGAGCTGGATATGAGCCAAAGCTTCA 647
QY 800 GAGCATGATGAGCGCGGAGTAAATGACAGCAATTTGGAAAAGATTACATCCAGCTGA 859
DB 648 GGAAGAAGTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 707
QY 860 AGAGCTGTTGACGCCATCCGATTTATTCAGAGCAGCTTATATGAGCCCTAATATG 919
DB 708 GCGAGCTCTGAGACGACACCTCTGACGACAGCTGACGCTGACGCTGACGCTGAC 767
QY 920 GCGCGCGAGGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 979
DB 768 GCGAGCGCGGAGGAGCAACCAACCACTGCTCAGAAAGCTTCAATGATGAGGAGGAA 827
QY 980 CAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1039
DB 828 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 887
QY 1040 ACTTCTGAAAACCTGCTGTTAGACACACTCTCTGACCAAGATTAGAAAATTCAGAA 1099
DB 888 ATTCTGAGACCCGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 947
QY 1100 TGTGTTATACATACATCCCAAGAAAGATTTGGCTTGAAGGTGATGATGATGATG 1159
DB 948 TGTGTTATACATACATCCCAAGAAAGATTTGGCTTGAAGGTGATGATGATGATGATG 1007
QY 1160 CTGAGGAGCAAAACATCTATCCGATTTCTATGCTGAGAGATTTCTATGTTGAACAT 1219
DB 1008 GCGGCGGCGGCGCGCGCTCTCAACACCTATGAGCGGCTTCAATGAGCTGAGCAAG 1067
QY 1220 TAGGAATGCTGAGCAATCAAGGCAATGATGATGATGATGATGATGATGATGATG 1279
DB 1068 TGGAGTTGGGCGCTGCGCGCTGAGCAATGATGATGATGATGATGATGATGATGATG 1127
QY 1280 GATACAACTACCTGTTGAGCAACAGATTTTAACCATTAACAGACTAGGCTCTCTCC 1339
DB 1128 GAGAGTATACCTGTTGAGCAACAGATTTTAACCATTAACAGACTAGGCTCTCTCC 1187
QY 1340 TCTACAAAGCGCTGATGAGCGCGCAAAAGCTTTGAGCTGATGATGATGATGATGATG 1399
DB 1188 TATACAAAGCGCTGATGAGCGCGCAAAAGCTTTGAGCTGATGATGATGATGATGATG 1247
QY 1400 AGCAGCGCTGAGCGCGCAAGTATGAGCGCGCAAAAGCTTTGAGCTGATGATGATGATG 1459
DB 1248 GAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1280
QY 1460 ACCAACAACAACACTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1519

```

```

DB 1281 CCCGCGACCCCAAAATACCGGAAAGGAGATGACACTGTTTGGCTTGAACCTTCCAA 1340
QY 1520 CAAGAAAGAAATCAAGCTGCTGGAGACTCTCAGAGCAAGCTGTTCAACAGTACTTCG 1579
DB 1341 TGACCCAGAGCTTGCAGCTGCTTGAAGAGTGTGAGTAAAGATGTGATCAGTACTCG 1400
QY 1580 TGCAAGCTTATGAGCGAGGAGGCTTAAAGTCCAGTACAGTACAGTACAGTACAGTAC 1639
DB 1401 TGCTGCGCCAGCGCAAGAGCAGATCTGTCCAGAGAGGTGACAGTGAATGGCGGCTAC 1460
QY 1640 TAGTATGATGAGAGCAGAGGAGGCTTCCAGATTTGAAGCCCGGCTTGGGCGGAGC 1699
DB 1461 TGCAATGATGAGAGAGTGAAGTACTCCCGCGCTGACAGAGATGAGGCTTGGCGGAG 1520
QY 1700 GGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1759
DB 1521 GAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1580
QY 1760 TGGGCTGC 1767
DB 1581 TTGCTTGC 1588

```

RESULT 14  
 US-10-431-438-10  
 ; Sequence 10, Application US/10431438  
 ; Publication NO. US20030180788A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: goldshmidt, orit  
 ; APPLICANT: pecker, itis  
 ; APPLICANT: vlodavsky, israel  
 ; APPLICANT: israel, michal  
 ; TITLE OF INVENTION: AVIAN AND REPTILE DERIVED POLYNUCLEOTIDE ENCODING A POLYPEPTIDE H  
 ; FILE REFERENCE: 26013  
 ; CURRENT APPLICATION NUMBER: US/10/431,438  
 ; NUMBER OF FILING DATES: 2003-05-08  
 ; NUMBER OF SEQ ID NOS: 16  
 ; SOFTWARE: Patentin version 3.1  
 ; SEQ ID NO 10  
 ; LENGTH: 1605  
 ; TYPE: DNA  
 ; ORGANISM: Gallus gallus  
 US-10-431-438-10

Query Match 16.5%; Score 292.8; DB 15; Length 1605;  
 Best Local Similarity 53.7%; Pred. No. 9.7e-86;  
 Matches 649; Conservative 0; Mismatches 532; Indels 27; Gaps 1;

```

QY 560 TTCTTTAAAGGAGCAATTCCTCAATCTTACGATATCTCATATTAAGCAGAGTCTC 619
DB 408 TGTCTCTCGCTGAACATCTCTGAAAAAGACAAAACACCAATTAACAAGAGCAGC 467
QY 620 TAGACAACTTTAACTTTGCTGATTTGCTGACCTCACTGATATTTGCTTAATG 679
DB 468 TGAACATCTCCACAGGTTGCGACAGAGCTCAGGCTTCGGCTGGATTTGGGCTGAAG 527
QY 680 CACTGCTGCTAATCCCAATTAATCTCTGAAACGTTTATGTCCTGAGTCTGTAAGT 739
DB 528 CACTGCTGCTGACAGGCTGAGCTGAGTGGAGACGCTTCAACGCAAGAGCTGAGGCT 587
QY 740 ACAGCGCCAGCAAAAGTATCAACATTTCTTGAGAACTGGATATGAGCCAAATTAATC 799
DB 588 ACTGTGACAGCGCAGCTCAACATCTCTGAGAGCTGGATATGAGCCAAAGCTTCA 647
QY 800 GAGCATGATGAGCGCGGAGTAAATGACAGCAATTTGGAAAAGATTACATCCAGCTGA 859
DB 648 GGAAGAAGTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 707
QY 860 AGAGCTGTTGACGCCATCCGATTTATTCAGAGCAGCTTATATGAGCCCTAATATG 919
DB 708 GCGAGCTCTGAGACGACACCTCTGACGACAGCTGACGCTGACGCTGACGCTGAC 767

```

Fri May 14 13:01:54 2004

us-10-088-676-1.rnpb

Page 13

QY	920	GGCGGCGGAGAAAGAAATGTCATGCGCCCTCCATAGATGATTCATGAAGGTGCGAGAAAGTA	979
Db	768	GGCGAGCCCGCAGACACACCGACCTGCTCAAGACCTTATGAATCTGGAGGGAAG	827
QY	980	CAGTAGATGCAGTTTACCTGGCAACATGTCATTTAGTGGCCGGGTGTCAGGTATGG	1039
Db	828	CGATTGACTCGGTCACCTGGCACCACTACTATGTGAATGAGCCGAAGTGCACAGAGGAGG	887
QY	1040	ACTTCCCTGAAAACTCGCGCTGTTAGACACCTCTCGACCAATATAGAAAATTCAGAAAG	1099
Db	888	ATTTCCTAGCCCTGAAGTGTGACCTCTTGGCATTGCCATACAGATGTCCTGGGGA	947
QY	1100	TGCTTAATACATACACTCCAGAGAAAGAAAGATTGGCTTGAAGGTGTGTGACCACTCAG	1159
Db	948	TCGTGAAGACAAAGGTGCGCCGCAAGAAAGATATGGCTGGGTGAGACCGGCTCGCTACG	1007
QY	1160	CTGAGGCGACAAACATCTATTCGCAATTCCTATGCTGCGAGAAATCTTATNGTGTGAACATT	1219
Db	1008	GCGGGGGGGGGCCCCCAGCTCTCCACACCTATGTGGCCGGCTCTCATGTGGCTGGAACAAGC	1067
QY	1220	TAGGAATGCTGGCCCAATCAGGCGCATTTGATGTGTCATACGGCAGCTCATTTTTCGACATG	1279
Db	1068	TGGGTTTGGCGCTCGGCGTGGCATTTGAATGTGTATGAGAGAGGTCTCTTTGGTGTG	1127
QY	1280	GATACATACCTCGTGGACCAAAATTTTAACCACTAACAGACTACTGGCTCTCTTCC	1339
Db	1128	GCAGCTATACCTGGTGGATGCGGCGCTTCAAGCCCTTCGCGACCTACGCGCTGTACCTGC	1187
QY	1340	TCTACAAAGCGCTGATCGGCGCCCAAAAGTCTTGGCTGTGATGTGGCTGGGCTCCAGCGGA	1399
Db	1188	TATACAAAGAGCTGTGTGGGCACCCGGGTACTACAGCCAGGCTGAGACAGAGGATGGGC	1247
QY	1400	AGCCACGCGCTGGCCGAGTGAATCCGGGACAACTAAGATTTTATGCTCACTGACCAAAACC	1459
Db	1248	GGCGCGCGC-----GGGTCTACCTGACATGCAACCAACC	1280
QY	1460	ACCACAAACCAACTAAGTCTGCGGTCCATCTACCTTTTATCATCAACTGGATCGAT	1519
Db	1281	CCCGGACCCCAAAATACCGGGGAAGGAGATGACACTGTTGGCTTGAACCTCTCCAAAG	1340
QY	1520	CAAGAAAGAAATCAGCTGCGCTGGGACTCTCAGAGACAAAGCTGATCCACAGTACTGCG	1579
Db	1341	TGACCCAGAGCTTGAAGCTGCTTAAGCAGTTGTGGAGTGAAGTGGATGAGTACGTACGCGC	1400
QY	1580	TGCAGCCCTTATGGGCAAGAGGCGCTAAAGTCCAAGTCAAGTCAACTGAATGGCCAGCCCT	1639
Db	1401	TGCTCTCCCAACGGCAAGGACAGCAATCTGTCTCAGAGAGGTGACGTGAATGGCCGCTTAC	1460
QY	1640	TAGTGAATGTGACAGACGGGACCTCTCCAGATTTGAAGCCCGCCCTTGGGGCGGCGC	1699
Db	1461	TGCAGATGTGACATGAGACACTTCCGCGCTCACAAGATGGGCCCTTGGCCCTGGGA	1520
QY	1700	GGACATTTGTCATCCTCCACAGTACCAATGGGCTTTTATGTGTCAAGAAATGTCAATGCTT	1759
Db	1521	GCACGCTCGGCTGCGACGCTTCTCTTAAGGTTTCTACGTGATCAGGAACGCTMAAGGCTTA	1580
QY	1760	TGGCGTCG 1767	
Db	1581	TTGCTTGC 1588	

```

RESULT 15
US-09-776-874A-43
; Sequence 43, Application US/09776874A
; Patent No. US20020102560A1
; GENERAL INFORMATION:
; APPLICANT: Becker, Iris
; APPLICANT: Vladavsky, Israel
; APPLICANT: Feinstein, Elena
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY
; TITLE OF INVENTION: EXPRESSION OF SAME IN GENETICALLY MODIFIED CELLS
; FILE REFERENCE: 01/22603
; CURRENT APPLICATION NUMBER: US/09/776, 874A

```

```

? CURRENT FILING DATE: 2001-12-12
? PRIOR APPLICATION NUMBER: US 08/922,170
? PRIOR FILING DATE: 1997-09-02
? PRIOR APPLICATION NUMBER: US 09/109,386
? PRIOR FILING DATE: 1998-07-10
? PRIOR APPLICATION NUMBER: PCT/US98/17954
? PRIOR FILING DATE: 1998-08-31
? NUMBER OF SEQ ID NOS: 47
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 43
? LENGTH: 2396
? TYPE: DNA
? ORGANISM: Homo sapiens
? US-09-776-874A-43

Query Match: 15.9%; Score 282.6; DB 9; Length 2396;
Best Local Similarity: 52.8%; Pred.No. 3,28-82;
Matches 677; Conservative 0; Mismatches 574; Indels 30; Gaps 2;

```

QY	474	AGATTAACAGAAAGGCTGCAGAAATGGCCAGACACCCGATGATATGCTGGCTCCAAAG	533
Db	932	AGTCAACCAAGATATTTGGAGGCTGAGCCGGTCTGTGCGGGTGTGAGAAACTCCA	991
QY	534	GGAAAGGACGCTCAATGCATCTGGTCTTTCTTAAAGGACCAATCTCCAACTACTTAAG	593
Db	992	GGTGAATATGGCCCTCCAGAGAGCTGGTGTGCTCCAGACACAGTACCAAAAGAGTTCTGA	1051
QY	594	TAATCTCATATTAACAGCCAGGCTCTAGACAACAACTTATTAACCTTGTGCTGTGG	653
Db	1052	GAAACGACCTACTCCAGAAAGCTCAGTGCACATGCTCTTACGTTTGTCCAAAGTCTCGGG	1111
QY	654	ACTCCACTGATATTTGCTCTTAAATGCATGCGTGTATATCCAAATAACTCTTGGAAACAG	713
Db	1112	GTTAGACCTGATCTTTGGTCTTAAATGGGTAACTACGAAACCCAGACTTAAGGTGAACAG	1171
QY	714	TTCTAAGGCCCTGAGCTGTGTGAATGATACAGGCCAGAGAAAAGTAAACAACATTTTGGGA	773
Db	1172	CTCCAGGCCAGCTTCTCTCTTGACTATGCTCTTCCAAAGGTTTAAACATCTCTCGGA	1231
QY	774	ACTGGGTATAGGCCAAATATCTATCGSAGCCATGCGCTGGCGAGTAAATGGACCCA	833
Db	1232	ACTGGGCATATAGCCCAACAGTTTCTGGAAGAAAGCTCACATCTCTCATGCTGGTGTGCA	1291
QY	834	GTTGGGAAAGATTCATCCAGCTGGAAGAGCTTTGACAGCCATCCGGAATTTATTCAG	893
Db	1292	GTTAGGAATAACCTTTGGAGGTGCATTAACCTTCTCA--AAAGTCAGTTTCCAAA	1348
QY	894	AGCCAGCTTAATATGAGCCCTAATATTTGGCGGCCAGAGAAAGATGATCATGCGCCTCTAGA	953
Db	1349	TGCAAAATCTATGCTCTGTGACATCGGTCAAGCTCGAGGGAAGACATTAACGTGTGAG	1408
QY	954	TGGAATTCATGAAGGTGCGAGGAAGTACAGTATGATGCAATTACTTCTGGCAACTTGGCTCAT	1013
Db	1409	GAGTTTCTTGAAAGGTGTGGCGAGAAAGTATGATGATCTCTTTACATGAGCACTCATTAATT	1468
QY	1014	TGATGCGCGGGGTGTCAGATGATGGACTTCCCTGAATACTGCGCTGTAGAACAACCTCTC	1073
Db	1469	GATATGAACGATCGTACCAAAAGAAAGATTTCTAGGCTCTATGGCGTGAACATTTTAT	1528
QY	1074	TGACCAATTAAGAAAATTCAGAAAAGTGTTAATACATCACTCCAGAAAAGAAAGTTTG	1133
Db	1529	TCTCTCTGTGCAAAAATTCGAAGGTCACTAAAGATCAACACTGGCAAGAAAGTCTG	1588
QY	1134	GCTTAAAGTGTGATGCCAATCCTGAGCTGAGGCGCAAAACAATCTATCCGATTCCTATGCG	1193
Db	1589	GTTGGAGAGAGACGACTCAGCTTACGATGGCGGTGCAACCTTGTGTCAACACTTTGCG	1646
QY	1194	TGCAGGATTCATTAGGTGAACATTTAGGAATGCTGGCAATCAGGCAATGATGTGTCT	1255
Db	1649	ACCTGACTTATATGTGCTGATTAATTTGGGCTGTCAACCCAGATGGGCATTAAGAGTGT	1708
QY	1254	GATTAGGCACATATTTTGTGACATGATCAAAATCACTCGGTGACCAAGATTTTAACCC	1311

Fr1 May 14 13:01:54 2004

us-10-088-676-1.rnpb

Page 14

Db	1709	GATBAGGCAAGTGTCTTCGGAGAGGCAACTACCACTTAGTGATGAAACTTTGAGCC	1768
Qy	1314	ATTACCGAGCTACTGGCTCTCTCTCTCTACTACAGCGCTGATCGGCCCCAAGCTTGGC	1373
Db	1769	TTTAACTGATTACTGGCTCTCTCTCTCTGTTCAAGAAACTGGTAGGTCCAGAGGTGTTACT	1828
Qy	1374	TGTGCTAGTGGCTGAGGCTCCAGACGAGCCACGAGCTGGCCGAGTGATCCCGGACAAACT	1433
Db	1829	GTCAAGAGTGAAAGGCCAGA-----CAGAGCAAACT	1861
Qy	1434	AAGATTTTATGCTACTGCTGACAAACCAACCAACCACTACGTTGCTGGTCCATTAC	1493
Db	1862	CCGAGTGATGTCACATGCACTTAACGTTATACCCACGAATACAGAGAGAGATCTAAC	1921
Qy	1494	ACTTTTATCATCAACTGTCATGCATCAAGAAAGAAATCAAGCTGAGCTGGACTCTCAG	1553
Db	1922	TCTGTATGTCCTGAACCTCCATATATGCAACAGACTTGAAGATACGGCTCCGTTGTT	1981
Qy	1554	AGACAGCTGTTCAACAGTACTGCTGACAGCCCTTATGGGCAAGAGGCTCTAAAGTCCA	1613
Db	1982	CAGGAAACCAAGTGGATGATGTAAGTCTTTGAAAGCTTCGGGCGCGAGTGAATTAATTCCTCA	2041
Qy	1614	GTCAGTGCACTGATGAGCGCAAGCCCTTAGTAGTGTGACAGCGGAGCCCTCCAGAAAT	1673
Db	2042	ATCTGTCCAACTGAAGCGTCAATTCTGAAGTGTGATGAGCAGAACCTTGCCAGCTTT	2101
Qy	1674	GAAAGCCCCGCCCCCTTGGGCGCGCGGACATGTGTAATCCCTCAGTCAACCATGGGCTT	1733
Db	2102	GACAGAAAAACCTCCCGCGAGGAAGTGCATAAGCCTGCCTGCTTTTCTGTATGCTTT	2161
Qy	1734	TTATGTGTCGAAGATGTCAA	1754
Db	2162	TTTTTCTAAAGAAATGCCAA	2182

Search completed: May 13, 2004, 17:38:08  
Job time : 796 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 13, 2004, 11:30:45 ; Search time 4726 Seconds  
(without alignments)  
11240.969 Million cell updates/sec

Title: US-10-088-676-1

Perfect score: 1779  
Sequence: 1 atgagagtgcttgcctt.....tggcctgcgcctaccgataa 1779

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database:

EST:  
1: em\_estda:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_estl:\*  
10: gb\_estc:\*  
11: gb\_hlc:\*  
12: gb\_estc:\*  
13: gb\_estc:\*  
14: gb\_estc:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vtl:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_man:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vtl:\*  
28: gb\_gssl:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	703.6	39.6	784	14	CF726388
2	505.4	28.4	10	BB116041	BB116041
3	420.2	23.6	12	BM426194	BM426194
4	387.4	21.8	620	BB627984	BB627984

Result	Score	Query Match	Length	DB ID	Description
5	306	17.2	573	12	BM083129
6	284.2	16.0	2173	11	AK040471
7	274.4	15.4	640	29	CE365946
8	258.8	14.5	351	9	A1019269
9	242	13.6	1962	11	AK087283
10	241.2	13.6	321	12	BI661209
11	232.6	12.0	838	13	BU474796
12	202.2	11.4	489	29	CG611539
13	191.6	10.8	1138	29	AL552151
14	187.6	10.5	334	29	CG653785
15	170.8	9.6	1200	9	AL545232
16	170.4	9.6	348	13	BX280673
17	166.6	9.4	427	29	AG019565
18	166.6	9.4	495	29	AG019564
19	162.4	9.1	777	9	AL718139
20	160	9.0	629	14	CB483444
21	150.6	8.5	631	14	CF915055
22	148.6	8.4	708	13	BQ775819
23	148.6	8.4	826	14	CK017201
24	144.6	8.1	433	9	AL917968
25	144.4	8.1	1083	13	BX398409
26	143.6	8.1	914	13	BX373611
27	139.6	7.8	285	29	CG486333
28	137.8	7.7	576	12	BQ037157
29	135.6	7.6	739	13	BU452690
30	130.2	7.3	907	13	BQ438834
31	126	7.1	652	13	BX107633
32	118.4	6.7	378	9	AI222323
33	118.2	6.6	682	12	BM956417
34	116.6	6.6	480	10	AM465382
35	114.2	6.4	556	9	AI342512
36	111.4	6.3	549	10	BF197674
37	110.8	6.2	292	29	CG511844
38	110.6	6.2	540	9	AI582254
39	110.5	6.2	773	14	CA387992
40	108.2	6.1	1185	9	AL552174
41	108.2	6.1	1201	9	AL545270
42	103.8	5.8	587	14	N45367
43	100.2	5.6	654	28	BH074791
44	99.2	5.6	445	10	BF040685
45	97.8	5.5	431	9	AA177901

## ALIGNMENTS

RESULT 1  
LOCUS CF726388 784 bp mRNA linear EST 09-OCT-2003  
DEFINITION UI-M-HBO-ckh-m-15-0-UI-r1 NIH\_BMAP\_HBO Mus musculus cDNA clone  
IMAGE:30547790 5', mRNA sequence.  
CF726388  
ACCESSION CF726388.1 GI:37600556  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

Mus musculus (house mouse)  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Struhsberg, Ph.D.  
Email: cga@bbs-remail.nih.gov  
Tissue Procurement: Dr. James Lin University of Iowa  
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Distribution information can be found at  
http://genome.uiowa.edu/distribution/mouse1.html  
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

The following repetitive elements were found in this cDNA

sequence: 280-300, >GC\_rich#low\_complexity  
 Seq primer: pyx-5,  
 Location/Qualifiers  
 source

1. 784  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:30547790"  
 /isvse\_type="whole eye"  
 /dev\_stage="embryo 12.5, 13.5, 14.5 dpc"  
 /lab\_pos="DH10B (T1 phage resistant)"  
 /clone\_lib="NH\_BMAP\_H80"  
 /note="Organ: Eye; Vector: pyx-Asc; Site: 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pyx-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TATGAGAGT. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH)."

## ORIGIN

Query Match 39.6%; Score 703.6; DB 14; Length 784;  
 Best Local Similarity 94.1%; Pred. No. 4.9e-192;  
 Matches 738; Conservative 0; Mismatches 45; Indels 1; Gaps 1;

122 CTGGAGACAGAGACCTTCCCTGCTGACAGAGCTGAGCTTGAAGAGAAAGACCTTA 181  
 1 CTGGAAACAGAGACCTTCCCTGCTGACAGAGCTGAGCTTGAAGAGAAAGACCTTA 60  
 182 TTCTACTGATGAGACCAAGAACCCAGTCAAGAGCTGAGCTTGAAGAGCTTCTCTC 241  
 61 TTCTACTGATGAGACCAAGAACCCAGTCAAGAGCTGAGCTTGAAGAGCTTCTCTC 120  
 242 TCGAGCTGATCCGTCATCATTCATGATGCTGCTGATTTCTTAACTCCAGCGCT 301  
 121 TCGAGCTGATCCGTCATCATTCATGATGCTGCTGATTTCTTAACTCCAGCGCT 180  
 302 TGGTGAACCTGGCCGGGAGCTTCCGCGCTTTCGCGCTTCGGGGGCAAGAGACCG 361  
 181 TGGTGAACCTGGCCGGGAGCTTCCGCGCTTTCGCGCTTTCGCGCTTCGGGGGCAAGAGACCG 240  
 362 ACTTCTGACATTCAGAACCTTGAAGAACCCGCGGAGCGCGCGCGCGCGCGCG 421  
 241 ATTTCCTGACATTCAGAACCTTGAAGAACCCGCGGAGCGCGCGCGCGCGCGCG 300  
 422 ATTACTATCAAAACTATGAGAGATGATGTTGGAAGTGAATGTTGCTTGAATTAAC 481  
 301 ATTACTATCAAAACTATGAGAGATGATGTTGGAAGTGAATGTTGCTTGAATTAAC 360  
 482 AGAAGGCTGCAAGATTGCCAGACCTGATGTTATGCTGCTGCTCAAGAGAGAG 541  
 361 AGAAGGCTGCAAGATTGCCAGACCTGATGTTATGCTGCTGCTCAAGAGAGAG 420  
 542 CAGCTGAGATGATGCTGCTTCTTAAAGAGCAATTCGCAATCTTCAAGATCTCA 601  
 421 CAGCTGAGATGATGCTGCTTCTTAAAGAGCAATTCGCAATCTTCAAGATCTCA 480  
 602 TATTAACAGCAGGCTCTTGAACAACCTTAACTTGAATGCTGCTGAGCTCAC 661  
 481 TATTAACAGCAGGCTCTTGAACAACCTTAACTTGAATGCTGCTGAGCTCAC 540  
 662 TGAATTTGCTTAAAGCACTGCTGCTAATCCCAATTAACCTCTGGAAGAGTTAG 721  
 541 TGAATTTGCTTAAAGCACTGCTGCTAATCCCAATTAACCTCTGGAAGAGTTAG 600

QY 722 CCTGAGCTGTTGAAGTACAGCGCCAGCAAAAGTACAACTTTCTGAGAACTGGGTA 781  
 DB 601 CCTGAGCTGTTGAAGTACAGCGCCAGCAAAAGTACAACTTTCTGAGAACTGGGTA 660  
 QY 782 ATGAGCCAAATTAATCTAGGACCATGATGCGCGGCGAGTAATGAGAGCGAGTGGAA 841  
 DB 661 ATGAGCCAAATTAATCTAGGACCATGATGCGCGGCGAGTAATGAGAGCGAGTGGAA 720  
 QY 842 AGATTATCAATCCAGCTGAGAGAGCTGTTGAGCCCATCCGATTTATTCAGAGCCAGCT 901  
 DB 721 AGA-TTATCCAGCTGATAGAGCGCTGCTCCANCCCATCCGAGTTTACTCAGAGCCAGCC 779  
 QY 902 TATA 905  
 DB 780 TGTG 783

RESULT 2  
 LOCUS BB116041  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 628)  
 Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Kono,M., Koyda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,Y., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,P., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.  
 RIKEN Mouse ESTs (Arakawa,T., et al. 2001)  
 Unpublished (2001)  
 On Jun 27, 2000 this sequence version replaced gi:8768609.  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsr.riken.go.jp/  
 URL: http://genome.gsc.riken.go.jp/  
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,M., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
 Wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Iwawa,M., Ohara,E., Watanabe,M., Yoneda,Y., Ishikawa,T., Oawa,K., Tanaka,T., Matsuda,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.  
 RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
 Kono,M., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.  
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Yamanaka,I., Kiyosawa,H., Kondo,S., Saito,T., Shinagawa,A., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K., Arakawa,T., Ishii,Y. and Hayashizaki,Y.  
 Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct. Funct. Genomics 2 pre, L72-L86 (2001)  
 Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.





Db	151	CCGGCCAAAGCCGCGGCGGGGCCCGGCCCGGATTACTACCTCAAGAACTACGAGAGCAGC	210
Qy	451	ATTGTTCCAAAGTGATGTTGCTTATGATTAACAGAAAGGCTCGCAAGATTGCCACGACCTT	510
Db	211	ATTGTTCCGAGGAGCAATTGCCCTGGATTAAAGAGAAAGCTGTAAAGATCGCCACGACATCCT	270
Qy	511	GATGTTATCTGATGCTCCAAAGGGAGAGGACAGTCAGATGCAATCGTCTCTTCAAG	570
Db	271	GACATTAATTTGGACCTACAGAGGAAAGGCACTCAGATGCACTTGTTGTTTAAAG	330
Qy	571	GAGCAATCTCCATATCTTACAGTATATCAATTTAAAGCGCAGCTCTTGAAGCAACTT	630
Db	331	GAGCAGTTTTCAAACACATACAGCAACTCACTTAACAGCAGGCTCTGACGAAACCTC	390
Qy	631	TATTAATTGTCGATGTCGTGGAACTCCACCTGATATTTGCTCAATATGACATGGGTGT	690
Db	391	TATTAATTGTGCTGACTGCTGTGCTTAACTGATCTTTGACTGAATGCTCTGCTGCA	450
Qy	691	AATCCCAATATCTCCTGGAAAGTTCTATGTCCTCGATCTGTGTGAAGTACAGCGCAAGC	750
Db	451	AATCCCAATATCTCCTGGAAAGCTCCAAATCGCTCAGCCTGGTGAAGTACAGTGCACGC	510
Qy	751	AAAAAGTACAACTTTCTTGGAAATGGGTATGAGCCAAATATCTATGGACATGAT	810
Db	511	AAGAAATTAACAATCTCCTCGGAGGTGTGGCAACAGGCCAATATCAACCGGATATGATT	570
Qy	811	GGCGGGGCAGTAAATGCGACCAAGTTTGGAAAGGATTAACATCCAGCTGAAGAG	863
Db	571	GATCGCTCAGTGAACGGCAGCCAGCTGGGGAAGGACTTAACATCCAGCTGAAGAG	623

LOCUS	BB621984	620 bp	mRNA	linear	EST 31-ANG-2001
DEFINITION	BB621984 RIKEN full-length enriched, adult male urinary bladder muscle.				
ACCESSION	musculus cDNA clone 953005C05 5', mRNA sequence.				
VERSION	BB621984				
KEYWORDS	BB621984.1 GI:15399264				
SOURCE	EST.				
ORGANISM	Mus musculus (house mouse)				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciuromorphi; Muridae; Murinae; Mus. 1 (bases 1 to 620)				
AUTHORS	Araoka,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Komio,H., Kondo,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shingawa,A., Shiraki,T., Sogabe,H., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeeda,Y., Tanaka,T., Toyota,T., Muramatsu,M., and Hayashizaki,Y. RIKEN Mouse ESTs (Araoka,T., et al. 2001)				
TITLE	Unpublished (2001)				
JOURNAL	Contact: Yoshihide Hayashizaki				
COMMENT	Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-72 Suenho-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-researc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/ Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Komio,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) wagci,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Matsubiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsurus,S., Kawai,T., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multipipillary sequencer. Genome Res.				

10 (Kono, H., 1757-1771 (2000))  
 10 (Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugihara, Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 261-269 (2001)  
 Yamanaka, I., Kiyosawa, H., Kondou, S., Saito, T., Shinagawa, A., Aizawa, K., Ishida, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., Arakawa, T., Ishii, Y. and Hayashizaki, Y.  
 Mapping of 19032 mouse cDNAs on mouse chromosomes. *J. Struct. Funct. Genomics* 2 pre, 172-186 (2001)  
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.  
 e mouse tissues:  
 Location/Qualifiers  
 1. 620

GAGGAGGAAGAAGTCACAGAGCTTTTTTTTTTTTTTTTNN 3'). cDNA was prepared by using klenowase thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trisepher. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 370.0. Second strand cDNA was prepared with the primer adapter of sequence 15' GAGGAGAGATTCGATTAAATAAATCCCCCCCCCCC 3'. cDNA was cleaved with XhoI and BamHI. Vector: a modified plasmid KS(+/-) after bulk excision from Lambda FLC I."

Query Match	21.8%	Score 387.4	DB 1.0	Length 620
Best Local Similarity	75.3%	Pred. No. 2e-100		
Matches 558	Conservative	0	Mismatches 31	Indels 162
				Gaps 2
QY	59	CGTGCCTAGCCCGGAGGCTCTCTACTTGGCTCTGTGTGTCATCTTCCCTTCTCC	118	
Db	2	CGTGCCTAGCCCGTGGCTCTCTTCTTGGCTCTGTGTGTCATCTTCCCTCTCTCC	61	
QY	119	AAGCTGAGAACAGAGAACCTTGTGCTGTATACAGAACTGCAGATTGGAAGAAAAGACC	178	
Db	62	AGCGTGAAGAACAGAGACCTTTGCTGTATACAGAACTGCAGATTGGAAGAAAAGACC	121	
QY	179	TGATTCCTACTGATGTGTAGACCAAGAACCCAGTCAGGACAGTCATGAGAACTCCTCT	238	
Db	122	TGATTCCTACTGATGTGTAGACCAAGAACCCAGTCAGGACAGTCATGAGAACTCCTCT	181	
QY	239	CTCTCAGCTGGAATCCGTCACTCATTCATGATGGCTGGCTCGATTTCTTAAGCTCCAAGC	298	
Db	182	CTCTCAGCTGGAATCCGTCACTCATTCATGATGGCTGGCTCGATTTCTTA-----	232	
QY	299	GCTTGATACCTCGGCCCGGAGACTTCCGCCCTTTCTCGGCTTCGGGGGCAAAAGGA	358	
Db	233	-----	232	
QY	359	CCGACTTCTCTGCAATTCAGAACCTTAGAGAACCCGGCGGAAAAGCCGCGGGGGCC	418	
Db	233	-----	232	

QY 419 CGGATTAATCTCTCAAAAATGAGATGACATTTGTCGAGATGATGTCCTTAGATA 478  
 DB 233 -----GACATTTGTCGAGATGATGTCCTTAGACA 263  
 QY 479 AACAGAAAGGCTGCAAGATGTCGCAACCTGATGATGATGTCCTGCAAGAGAGA 538  
 DB 264 AGCAAAAGGCTGTAGATGTCGCAACCTGATGATGATGTCCTGCAAGAGAGA 323  
 QY 539 AGGCAAGTCAAGATGATGTCGTCCTGCAAGAGATGTCCTGCAAGATGATGTC 598  
 DB 324 AGGCAAGTCAAGATGATGTCGTCCTGCAAGAGATGTCCTGCAAGATGATGTC 383  
 QY 599 TCATATTAACAGCAGGCTCTGACAACTTTATTAATCTTCTGATGTCCTGACATCC 658  
 DB 384 TCATATTAACAGCAGGCTCTGACAACTTTATTAATCTTCTGATGTCCTGACATCC 439  
 QY 659 ACTGATATTTGTCCTGCAAGATGTCGTCCTGATGTCCTGATGTCCTGATGTCCT 718  
 DB 440 ACTGATATTTGTCCTGCAAGATGTCGTCCTGATGTCCTGATGTCCTGATGTCCT 499  
 QY 719 GTGCTGATGTCCTGTCGCAAGATGTCGTCCTGATGTCCTGATGTCCTGATGTCCT 778  
 DB 500 GTGCTGATGTCCTGTCGCAAGATGTCGTCCTGATGTCCTGATGTCCTGATGTCCT 559  
 QY 779 GTATATGAGCAAAATTAATCTGAGATGTCGTCCTGATGTCCTGATGTCCTGATGTCCT 838  
 DB 560 GTATATGAGCAAAATTAATCTGAGATGTCGTCCTGATGTCCTGATGTCCTGATGTCCT 619  
 QY 839 G 839  
 DB 620 G 620

RESULT 5  
 BM083129 573 bp mRNA linear EST 14-NOV-2001  
 LOCUS fu29f07.y1 Campbell zebrafish ovary Danio rerio cDNA clone  
 DEFINITION IMAGE:5306653 5' similar to TR:Q9QZ8 Q9QZ8 HEPARANASE.; mRNA  
 sequence.  
 accession EM083129 GI:16930059  
 VERSION EST.  
 KEYWORDS Danio rerio (zebrafish)  
 SOURCE Danio rerio  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
 Cypriniformes; Cyprinidae; Danio.  
 1 (bases 1 to 573)  
 Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M.,  
 Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,  
 Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,  
 Peterson, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R.,  
 Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,  
 Waterston, R. and Wilson, R.  
 WashU Zebrafish EST Project 1998  
 Unpublished (1998)  
 TITLE TITLE  
 JOURNAL  
 COMMENT Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: zbrfish@watson.wustl.edu  
 CDNA library constructed by library constructed by Invitrogen and  
 donated by R. Campbell (Marine Biology Laboratory, Woods Hole, MA).  
 DNA sequencing by: Washington University Genome Sequencing Center  
 clone distribution: the I.M.A.G.E. Consortium/LML, send email to:  
 infoimage.lml.gov  
 High quality sequence stop: 439.  
 Location/Qualifiers  
 1..573  
 /organism="Danio rerio"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7955"

ORIGIN  
 Query Match 17.28; Score 306; DB 12; Length 573;  
 Best local similarity 72.8%; Pred. No. 7.5e-77;  
 Matches 409; Conservative 0; Mismatches 150; Indels 3; Gaps 1;  
 /clone="IMAGE:5306653"  
 /sex="female"  
 /dev stage="4-5 months, 1 year and 2 years"  
 /lab host="DH10B (phage-resistant)"  
 /clone lib="Campbell zebrafish ovary"  
 /note="Organ: ovary (pooled); Vector: pCMV-Sport6; Site: 1;  
 NotI; Site 2: SalI; Whole ovaries collected from zebrafish  
 aged 4-5 months, 1 year and 2 years. Oligo-dT primed,  
 directionally cloned. Average insert size 2 kb. Library  
 constructed by Invitrogen and donated by R. Campbell  
 (Marine Biology Laboratory, Woods Hole, MA)."

RESULT 6  
 AK040471 2173 bp mRNA linear HTC 19-SEP-2003  
 LOCUS Mus musculus 0 day neonate thymus cDNA, RIKEN full-length enriched  
 DEFINITION library, clone:A43010M04 product:heparanase, full insert sequence.  
 accession AK040471 GI:26333764  
 VERSION AK040471.1  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1  
 Carmichael, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning

QY 289 AGCTCCAGGCGCTTGTGACCTGGCCCGGAGACTTTCGCGCTTTCGCGCTTCGCGG 348  
 DB 4 ATCTCCAGGCGCTTGTGACCTGGCCCGGAGACTTTCGCGCTTTCGCGCTTCGCGG 63  
 QY 349 GGCAGAAAGGACCGACTTCTGCAAGTTCGAGACCTGAGAGAACCCGCGAGAGCGG 408  
 DB 64 GGCAGAAAGGACCGACTTCTGCAAGTTCGAGACCTGAGAGAACCCGCGAGAGCGG 123  
 QY 409 GGGCCCGGCGCGGATTCATCTCAAAAAGTATGAGAGAGAGAGAGAGAGAGAGAG 468  
 DB 124 ---CTGACCGAGCTATTAAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
 QY 469 GCCTTATTAACAGAGAGGCTGCAAGATGTCGAGACCTGAGATGTCGAGATGTCG 528  
 DB 181 GCTTGTGAGAGAGAGAGAGGCTGAGATGTCGAGACCTGAGATGTCGAGATGTCGAG 240  
 QY 529 CAAAGGAG 588  
 DB 241 CAGAGGAG 300  
 QY 589 TACAGTATGCTCATTAATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 648  
 DB 301 TACAGTATGCTCATTAATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
 QY 649 TCTGAGTTCACCTGATATTTGCTTAATTAACAGAGAGAGAGAGAGAGAGAGAGAG 708  
 DB 361 GCAGAGTTCACCTGATATTTGCTTAATTAACAGAGAGAGAGAGAGAGAGAGAGAG 420  
 QY 709 AACAGTTCAGTTCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 768  
 DB 421 AATGCAACAGTTCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
 QY 769 TGGAGACTGGGTAATGAG 828  
 DB 481 TGGAGACTGGGTAATGAG 540  
 QY 829 AGCGAGTTGGAG 850  
 DB 541 TCTGCTGCTGCTGAG 562



Db 1086 GTTGGAGAGAGAGAGCTAGCTTAAAGGTGGCGGTGACACCTTGTCTGTCACACCTTTGC 1145  
 QY 1194 TGCAGATTCTTATGTTGTAACAATTAGGAATGCTGGCCAAATCAGGGCTATGATGTGT 1253  
 Db 1146 AGCTGGCTTATGAGCTGATTAATTTGGGCTGTGACGCCAGATGGGCTAGAAAGTGT 1205  
 QY 1254 GATACGGCACTATTTTGTGACATATACATCACTGCTGAGACAGAAATTTAAACC 1313  
 Db 1206 GATGAGCAGAGTCTTCTGAGAGAGCACTACCTAGTGTGTAATTTGAGCC 1265  
 QY 1314 ATTACAGACTACTGCTCTCTCTCTCTCAACAAGCCCTGATGAGCCCAAAAGTCTGAGC 1373  
 Db 1266 TTATCTGATGATGCTCTCTCTCTCTCTCAAGAACTGATGATCCAGGCTGTACT 1325  
 QY 1374 TGTGATGATGCTGAGGCTTCAAGGGAAGCAGGCTGAGCTGAGTATCCGGGAAACT 1433  
 Db 1326 GTCAAGAGTAAAGGCCAGAG-----CAGAGCAACT 1358  
 QY 1434 AAGGATTATGCTCACTGCAACAACCAACAACAACAACAACAACAACAACAACAACAACA 1493  
 Db 1359 CCGAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1418  
 QY 1494 ACTTTTATATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1553  
 Db 1419 TGTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1478  
 QY 1554 AGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1613  
 Db 1479 CAGAGAAACAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1538  
 QY 1614 GTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1673  
 Db 1539 ATCTGCTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1598  
 QY 1674 GAGAGCCCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1733  
 Db 1599 GAGAGAAACCT 1658  
 QY 1734 TTATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1754  
 Db 1659 TTTGCTATGAGAAATGCCAA 1679

RESULT 7  
 CE365946 640 bp DNA linear GSS 27-SEP-2003  
 LOCUS tigr-gss-dog-17000362155477 Dog Library Canis familiaris genomic.  
 DEFINITION genomic survey sequence.  
 ACCESSION CE365946  
 VERSION CE365946.1 GI:36583778  
 KEYWORDS GSS.  
 SOURCE Canis familiaris (dog)  
 ORGANISM Canis familiaris

REFERENCE AUTHORS  
 TITLE The dog genome: survey sequencing and comparative analysis  
 JOURNAL Science 301 (5641), 1898-1903 (2003)  
 MEDLINE 22875432  
 PUBMED 14512627  
 COMMENT Contact: Kirkness EF  
 The Institute for Genomic Research  
 Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,  
 Rockville, MD 20850, USA  
 Tel: 301-838-0200  
 Fax: 301-838-0208  
 Email: ekirkness@tigr.org  
 Class: shotgun.  
 Location/Qualifiers  
 1..640

FEATURES  
 source

Query Match 15.4%; Score 274.4; DB 29; Length 640;  
 Best Local Similarity 96.2%; Pred. No. 1.2e-67;  
 Matches 281; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
 /organism="Canis familiaris"  
 /mol\_type="genomic DNA"  
 /strain="standard poolie"  
 /db\_xref="taxon:9615"  
 /clone\_1lb="Dog Library"  
 /note="Site 1: BstXI; Libraries were prepared from peripheral blood"

ORIGIN  
 1 ATGAGGCTGCTTTGCTTCCCTTCCCTGAGCCATGCTCCAGCAACTCCGCCCCCGCG 60  
 174 ATGAGGCTGCTTGTGCTTCCCTGAGCCATGCTCCAGCAACTCCGCCCCCGCG 233  
 61 TGCCTAGCCCCGGGGGCTCTCTACTTGGCTCTGTGTCTCATCTCTCCCTTCTCCAG 120  
 234 TGCCTAGCCCCGGGGGCTCTCTACTTGGCTCTGTGTCTCATCTCTCTCTCTCCAG 293  
 121 GCTGAGACAGAGACCTTGGCTGTAGACAGAGCTGACAGTTGAGAGAAAGACCTG 180  
 294 GTTGAAGACAGAGACCTTGGCTGTAGACAGAGCTGACAGTTGAGAGAAAGACCTG 353  
 181 ATTCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240  
 354 ATTCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 413  
 241 CTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 292  
 414 CTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 465

RESULT 8  
 A1019269 351 bp mRNA linear EST 16-JUN-1998  
 LOCUS ub22a07.t1 Soares\_thymus\_2NbwT Mus musculus cDNA clone  
 DEFINITION IMAGE:1378452 5', mRNA sequence.  
 ACCESSION A1019269  
 VERSION A1019269.1 GI:3233605  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

REFERENCE AUTHORS  
 TITLE The WashU-HMNI Mouse EST Project  
 JOURNAL Unpublished (1996)  
 COMMENT Contact: Marra M/Mouse EST Project  
 WashU-HMNI Mouse EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@wustl.edu  
 This clone is available royalty-free through LBNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 MGI:900920  
 Seq primer: -28m3 rev2 ET from Amersham  
 High quality sequence stop: 337.  
 Location/Qualifiers  
 1..351  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:1378452"

FEATURES  
 source







Db	216	CTATGGGACGAAGCGCTTAAAGTTCCAGTCAGTGCACTCAATGGCCAGCCTTAGTGA	275
QY	1646	TGTGTGACGACGGGACCCTCCAGATTGAAGCCCCGCCCTTCG	1691
Db	276	TGTGTGACGACGGGACCCTTCAGAACTGAAGCCCCGCCCTGC	321
RESULT 11			
LOCUS	BU474796		
DEFINITION	838 bp mRNA linear EST 30-NOV-2002		
ACCESSION	603763360F01 CSEORBN21 Gallus gallus cDNA clone CHEST68516 5', mRNA sequence.		
VERSION	BU474796		
KEYWORDS	EST.		
SOURCE	BU474796.1 GI:25968373		
ORGANISM	Gallus gallus (chicken)		
REFERENCE	Gallus gallus		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.		
TITLE	1 (bases 1 to 838).		
JOURNAL-MEDLINE	Boardman,F.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E., Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J. A Comprehensive Collection of Chicken CDNAS Curr. Biol. 12 (22), 1965-1969 (2002) 22335534 12445392		
PUBMED	Contact: Simon Hubbard Department of Biomolecular Sciences University of Manchester Institute of Science and Technology (UMIST) PO Box 88, Manchester, M60 1QP, UK Tel.: 01612008930 Fax: 01612360409 Email: Simon.Hubbard@umist.ac.uk. Location/Qualifiers		
FEATURES			
Source	1..838		
	/organism="Gallus gallus"		
	/mol_type="mRNA"		
	/strain="layer"		
	/db_xref="taxon:9031"		
	/clone="CHEST68516"		
	/sex="female"		
	/dev_stage="adult"		
	/lab_host="DH10B"		
	/clone_idb="CSEORBN21"		
	/note="Organ: ovary; Vector: plasmidscript II KS(+); Site_1: EcoRI; Site_2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dt) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was bluntended, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the plasmidscript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."		
ORIGIN			
Query Match	12.0%; Score 212.6; DB 13; Length 838;		
Best Local Similarity	79.7%; Pred No. 1.2e-49;		
Matches 251; Conservative	0; Mismatches 64; Indels 0; Gaps 0;		
QY	1005	TTGTCTACATTGATGCGCGGATGTGTCACAGGTATGAGACTTCTGAAAGCTCGCTGTAGA	1064
Db	93	TTACTATCATCATGCGCCAGTGGCCAAAGTGACTGACTTCCGAAAAACACGCTGCTGGA	152
QY	1065	CACACTCTCTGACCAAGATTAGAAAATTCAGAAAGTGTATATACATCACCTCCAGGAA	112

Db	155	CACACTGTTCGACCAAGATCAGGAAAAATCCAGAAAGTTTGAAAGCATACACACCGGGGAA	212
Qy	1125	GAAGATTTGGCTTGAAGGTGTGTGTGACCACTCAGCTGAGGCAACAAACATCTATCCGA	1184
Db	213	GAAGATCTGGCTGGAAGGCGCTGGTGTCTACTCTCTGAGGCAAGAAACATCTTCCGA	272
Qy	1185	TTTCCATGCTGCAAGATTTCTTAAGTGTGAACAATTAGAAAGCTGGGCAATCAGGGCAT	1244
Db	273	TTTCCATGCTGCGCGGGTTCCTGTGGCTGAACAGCTGGGCTTGTGGCCAGCAAGGAT	332
Qy	1245	TGATGTGCTGATACAGCACTCATTTTTCGACCATGATCAATCACTCTGTGACCAAGA	1304
Db	333	TGATGTAGTGGTGCGGCACTTCTCTTCCTTGACCATGAGCAACACACTGTGTGACAGAA	392
Qy	1305	TTTTAACCATTACC	1319
Db	393	CTTCAACCCCTCGCC	407
RESULT 12			
LOCUS	CG611539		
DEFINITION	CG611539	483 bp	DNA
ACCESSION	CG611539		linear
VERSION	CG611539.1		GenSeq
KEYWORDS	SSS.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; 1 (bases 1 to 483)		
AUTHORS	Zambrowicz,B.P., Abulín,A., Ramirez-Solis,R., Richter,L.J., Pigott,J., BeltrandeRio,H., Buxton,E.C., Edwards,J., Finch,R.A., Fridde,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaiing,C., Key,B.W., Jr., Kild,P., Kohlhauff,B., Ma,Z.-Q., Markesich,D., Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z., Sparks,M.U., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N., Zhu,Q., Peterson,C. and Sands,A.T.		
TITLE	Mnk1 kinase deficiency lowers blood pressure in mice: a gene-trap screen to identify potential targets for therapeutic intervention		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	100 (24)	14109-14114 (2003)
COMMENT	Contact: Zambrowicz BP		
	OmitBank		
	Lexicon Genetics Incorporated		
	4000 Research Forest Drive, The Woodlands, TX 77381, USA		
	Email: material@lexgen.com		
	Gene trap sequence tag generated by 3' RACE from mouse ES cells as described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)		
	Class: Gene Trap		
FEATURES			
source	1..483		
	location/Qualifiers		
	/organism="Mus musculus"		
	/mol_type="genomic DNA"		
	/strain="129SV/EV"		
	/db_xref="taxon:10090"		
	/clone="OST296636"		
	/cell_type="embryonic stem cell"		
	/clone_id="Mus musculus 129SV/EV"		
ORIGIN			
Query Match	11.4%	Score 202.2	DB 29, Length 483
Best Local Similarity	87.0%	Pred. No. 8.3e-47	
Matches 241, Conservative	0	Mismatches 34	Indels 2, Gaps 2
Qy	1340	TTTACAAGGCTGATCGGCCCCCAAGCTCTTGGCTGTGCATGTGGCTGGGCTTCAGCGGA	1399
Db	1	TTTACAAGGCTGTGATTGGCCCCCAAACTCTGGCAGGACATGATGCTGGAGCTCAGCGGA	60
Qy	1400	AGCCACGGGCTGCGCGAGTATCCGGGACAAACTAAGATTATGCTCATCTCAGAAAC	1458
Db	61	AGCTCTGGCAGGAAGATATCCGAACAAACTTNGAATCTATGCTCATCTCAGAAAC	120

QY 1460 ACCACACCACTACGT-TCGTGGTCCATTACA-CTTTTATCATCACTTGCATCG 1517  
 DB 121 ACCACACCACTACGTTCCTGGTGGTTCATCAACCTTTTATCATCACTTGCATCG 180  
 QY 1518 ATCAAGAAAGAAATCAAGCTGGTGGTTCATCAAGCAAGCTGGTTCACCACTTCT 1577  
 DB 181 ATCAAGAAAGAAATCAAGCTGGTGGTTCATCAAGCAAGCTGGTTCACCACTTCT 240  
 QY 1578 GCTGACGCTTATGGGACAGAGGCTTAAGTCCAG 1614  
 DB 241 GCTGACGCTTATGGGACAGAGGCTTAAGTCCAG 277  
 RESULT 13  
 AL552151/c 1156 bp mRNA linear EST 31-MAY-2003  
 LOCUS AL552151 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA  
 DEFINITION Clone CS0D1059YN15 3-PRIME, mRNA sequence.  
 ACCESSION AL552151  
 VERSION AL552151  
 KEYWORDS GI:31273967  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 1 (bases 1 to 1156)  
 Mammalian, Eutherian, Primate, Catarrhini, Homiidae, Homo.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT On Feb 15, 2001 this sequence version replaced gi:12890775.  
 CONTACT: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: sequef@genoscope.cns.fr, Web: www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 2469.r For  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0D1059CG08NP1&cluster=2469.r. Contact :  
 Feng Liang, Email: fliang@life.techn.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS0D1059CG08NP1.  
 Location/Qualifiers  
 1..1156  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1059YN15"  
 /isue="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo (dt)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Query Match 10.8%; Score 191.6; DB 9; Length 1156;  
 Best Local Similarity 51.6%; Pred. No. 1.8e-43;  
 Matches 525; Conservative 3; Mismatches 458; Indels 31; Gaps 3;

QY 750 CAAAAGTACAAATTTCTTGGAGACTGGGTAATGAGCAAAATTAATTCGACCATGCA 809  
 DB 986 CCAAGGATMTAAATTTCTTGGAGACTGGGTAATGAGCAAAATTAATTCGACCATGCA 928  
 QY 810 TGGCCGGGACGATTAATGAGCAAGGTAATGAGCAAAATTAATTCGACCATGCA 869  
 DB 927 TGAATTTTATCAATGAGGTCGAGTGAAGAGATTTTATTCATTCATCAATTAAC---T 871  
 QY 870 GCAAGCCATCCGATTTATTCAGAGCCGCTTAATGAGGCTTAATTTGGGCGCGGAG 929  
 DB 870 TCTAAGAAAGTCCACCTTCGAAATTAATTAATTCATGAGGCTTAATTTGGGCGCGG 811  
 QY 930 GAGAAATGTCATCGCCCTCTTAATGAGATTCATGAAGTGGCAGAAAGTACAGTAGATGC 989

DB 810 AGAAAGACGGCTAAGATGCTGAAGAGCTTCTGAAGGCTGGGAGAGAGATGATGATTC 751  
 QY 990 AGTTACTGGCAACATTTGCTACATGATGATGAGCGGGGTCGACAGTATGAGACTTCTGTA 1049  
 DB 750 AGTTACTGGCAACATTTGCTACATGATGATGAGCGGGGTCGACAGTATGAGACTTCTGTA 691  
 QY 1050 AACTGCTGCTTGAAGACATCTCTGACAGATTAAGAAATTCAGAAAGTGTAAATAC 1109  
 DB 690 CCGTGATGATGAGACATTTTATTCATCTGCAAAAAGTTTCCAGTGGTGAAGAG 631  
 QY 1110 ATACATCTCAGAAAGAAAGTTTGGCTTGAAGGTGTGAGCACTCAGCTGGAGGAC 1169  
 DB 630 CACACGCTCGGCAAGAGGCTGTGTAAGAGAAACAGCTGCAATATGAGAGCGAGC 571  
 QY 1170 AAACATCTATCCGATTTCTATGCTGACAGATTTCTTAATGTTGAACATTTAAGAAATGCT 1229  
 DB 570 GCCCTTGCTATCCGACACCTTTGACAGCTGCTTATGAGCTGATTAATTTGGCCGCTC 511  
 QY 1230 GGCATATCAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1289  
 DB 510 AGCCGATGAGGAAATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 451  
 QY 1290 CCGTGGACAGCAATTTTAACCATTAACCACTGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1349  
 DB 450 TTTATGATGAGAAACTTGCATCTTATCTGATGATGATGATGATGATGATGATGATGATGATGAT 391  
 QY 1350 CCGTATGAGGCCCCAAGATCTTGGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1409  
 DB 390 ATGATGAGGACCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 339  
 QY 1410 TGGCCGAT 1469  
 DB 338 -----AAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 298  
 QY 1470 CAATGAT 1529  
 DB 297 AAGGATTAAGAAAGGATTTAATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 238  
 QY 1530 AATCAAGCTGCTGAGACTCTCAGAGCAGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1589  
 DB 237 CTTCGAT 178  
 QY 1590 TGGCCGAT 1649  
 DB 177 GGAAGCTGAT 118  
 QY 1650 GGAAGCTGAT 1709  
 DB 117 GGAAGCTGAT 58  
 QY 1710 CATCCCTGAT 1766  
 DB 57 CTTCGAT 1

## RESULT 14

CG653785 334 bp DNA linear GSS 02-OCT-2003  
 LOCUS CG653785  
 DEFINITION OST419746 Mus musculus 129Sv/Ev Mus musculus genomic clone  
 OST419746, genomic survey sequence.

ACCESSION CG653785  
 VERSION CG653785.1 GI:37477634  
 KEYWORDS GSS.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE Zammit, B.P., Abidin, A., Ramirez-Solis, R., Richter, L.J.,  
 Piggott, W., Beltrande-Rio, H., Buxton, E.C., Edwards, D., Finch, R.A.,  
 Fridde, C.J., Gupta, A., Hansen, G., Hu, Y., Huang, W., Jaling, C.,  
 Key, B.W. Jr., Kipp, P., Kohlhaufl, B., Ma, Z.-Q., Markesich, D.,

Payne, R., Potter, D.G., Qian, N., Shaw, J., Schrick, J., Shi, Z.-Z., Sparks, M.J., Van Sligtenhorst, I., Vogel, P., Walke, W., Xu, N., Zhu, Q., Person, C. and Sands, A.T.  
 Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap screen to identify potential targets for therapeutic intervention  
 Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)  
 Contact: Zambrowicz BP

**JOURNAL COMMENT**  
 Omnibank  
 Lexicon Genetics Incorporated  
 4000 Research Forest Drive, The Woodlands, TX 77381, USA  
 Email: materials@lexgen.com  
 Gene trap sequence tag generated by 3' RACE from mouse ES cells as described in Zambrowicz et al (Nature, 1998 Apr 9;392(6676):608-11)  
 Class: Gene Trap.  
**FEATURES**  
 source  
 Location/Qualifiers  
 1..334  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="129SV/ev"  
 /db\_xref="taxon:10090"  
 /clone="OST419746"  
 /cell\_type="embryonic stem cell"  
 /clone\_id="Mus musculus 129SV/ev"

## ORIGIN

Query Match 10.5%; Score 187.6; DB 29; Length 334;  
 Best Local Similarity 82.6%; Pred. No. 1,le-42;  
 Matches 242; Conservative 0; Mismatches 47; Indels 4; Gaps 3;

QY 1325 ACTGGCTCTCTCTCTTATCAAGCGCTGATCGCCCAAGTCTGGCTGTCATGTGG 1384  
 Db 1 ATGGCTCTCTCTCTCTTCAAGCGTGTATGGCCCAAGTCTGGCAAGTCAATGTG 59  
 QY 1385 C-TGGGCTCAGCGAAGCAAGCGCTGCGCGAGTATCGGCAAACTAAGATTTAT 1443  
 Db 60 CGGGGCTCAGCGAAGCGCTGCGCGAGTATCGGCAAACTAAGATTTAT 119  
 QY 1444 GCTCACTGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1501  
 Db 120 GCTCACTGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 179  
 QY 1502 TCATCACTGCAATGATCAAGAAAGAAATCAAGGCTGGGCTCTCAAGACCAAC 1561  
 Db 180 TCATCACTGCAATGATCAAGAAAGAAATCAAGGCTGGGCTCTCAAGACCAAC 239  
 QY 1562 TGGTTCAACCAATCTGCTGAGCCCTATGCGGAGAGGAGCTTAAGTCCAAAG 1614  
 Db 240 TTGTGNCACCAATGCTGAGCCCTATGCGGAGAGGAGCTTAAGTCCAAAG 292

**RESULT 15**  
 AL545232/c 1200 bp mRNA linear EST 31-MAY-2003

LOCUS AL545232 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 DEFINITION clone GS01028YF04 3-PRIME, mRNA sequence.  
 AL545232

ACCESSION AL545232 GI:31267068  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 1200)  
 Li, W.B., Gruber, C., Jessse, J. and Polyes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On Feb 15, 2001 this sequence version replaced gi:12877713.  
**REFERENCE**  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 2469.r For

more information about this cluster, see  
<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=GS01028DC02NP1&cluster=2469.r>. Contact :  
 Feng Liang Email : fliang@lifetech.com URL :  
<http://fulllength.invitrogen.com/InvitrogenCorporation1600FaradayAvenueGenoscopeSequenceID:GS01028DC02NP1>.

**FEATURES**  
 source  
 Location/Qualifiers  
 1..1200  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="GS01028YF04"  
 /issue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_id="Homo-sapiens PLACENTA COT 25-NORMALIZED"  
 /note="Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Query Match 9.6%; Score 170.8; DB 9; Length 1200;  
 Best Local Similarity 51.6%; Pred. No. 2e-37;  
 Matches 431; Conservative 3; Mismatches 375; Indels 27; Gaps 1;

QY 902 TATATGGCCCTATATTTGGGCGCGGAGAGATGTCATGCGCCCTCTAGATGATTTCA 961  
 Db 809 TCTATGGCTGATGTTGTCAGCTCGAAGAAAGAGCGMAAATGCTGAAGAGCTTCC 750  
 QY 962 TGAAGTGGCAGAGAGTACATGATGATGATGATGATGATGATGATGATGATGATG 1021  
 Db 749 TGAAGTGGTGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 650  
 QY 1022 GGGTGTCAAGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1081  
 Db 689 GGAATGCTACAGAGAGAGATTAATTAACCTGATGATGATGATGATGATGATGATG 630  
 QY 1082 TTAGAAAATTCAGAAAGTGGTTATATCATCATCATCATCATCATCATCATCATCAT 1141  
 Db 629 TGCAAAAAGTTTCAGAGTGGTGGAGACACAGCGCTGGCAAGAGTGGTGGAGAG 570  
 QY 1142 GTGTGTGACACCTGAGCTGAGGACCAACAATCATGATGATGATGATGATGATGAT 1201  
 Db 569 AAAGAAGCTGATGATGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 510  
 QY 1202 TCTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1261  
 Db 509 TTATGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 450  
 QY 1262 ACTCATTTTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1321  
 Db 449 AAGATATCTTTGAG 390  
 QY 1322 ACTATGCT 1381  
 Db 389 ATTATGCT 330  
 QY 1382 TGGCTGGGCTCAGCGAT 1441  
 Db 329 TGCAAGGTTCAAGAGAGAG-----AAGCTTCAGATAT 297  
 QY 1442 ATGCTCATGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1501  
 Db 296 ACCCTCATGTCAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 237  
 QY 1502 TCATCACTGCAATGATCAAGAAAGAAATCAAGGCTGGGCTCTCAAGACCAAG 1561  
 Db 236 CCAATTAACCTCATAATGATCAAGAGTCTGGGTTACCTTATCTTTTCTTACAGC 177  
 QY 1562 TGGTTCAACCAATCTGCTGAGCCCTATGCGGAGAGAGGCTTAAGTCCAAAGTCC 1621  
 Db 176 AAGTGATTAATATCTTCTAAGACCTTTGAGAGCTCACTGATTAATCTTCAAAATCT 117  
 QY 1622 AACTGAATGGCAGAGCCCTTATGATGATGATGATGATGATGATGATGATGATGAT 1681

Db 116 AACTCATTGCTTAAAGTGTGATATCAAACTTGGCACTTTAATGAAA 57  
QY 1682 GCCCCCTTCGAGCCGCGGACATTGCTCATCCCTCAGTCACCAATGAGCTTTAT 1737  
Db 56 AACCTCTCGGCGCAGGAAGTTCACTGGGCTTGCAGATTCTCATATATTTT 1

Search completed: May 13, 2004, 15:37:47  
Job time : 4747 secs







QY 421 HSPFDHGYNHLVDONENPLPDYWLSLYKRLIGPKYLAHVAGLQRPKPRPVIRDKLRI 480  
 CC culture of primary tissues, to regenerate tissues and in chemotaxis. The  
 CC polypeptides can also be used as a food additive or preservative to  
 CC increase or decrease storage capabilities. The present sequence  
 CC represents the amino acid sequence of human heparanase-like protein,  
 CC splice variant #1  
 Db 481 YACCTNHNHNHYVRSITLFTIINLHRSRKKIKLAGTRDKLVHGYLLQPYGQEGLSKSV 540  
 QY 481 YACCTNHNHNHYVRSITLFTIINLHRSRKKIKLAGTRDKLVHGYLLQPYGQEGLSKSV 540  
 Db 481 YACCTNHNHNHYVRSITLFTIINLHRSRKKIKLAGTRDKLVHGYLLQPYGQEGLSKSV 540  
 QY 541 QUNGQPLVMVDGTLPELKPRLRAGRTLVIPVTMGFYVKNVNALACRYR 592  
 Db 541 QUNGQPLVMVDGTLPELKPRLRAGRTLVIPVTMGFYVKNVNALACRYR 592  
 QY 541 QUNGQPLVMVDGTLPELKPRLRAGRTLVIPVTMGFYVKNVNALACRYR 592  
 Db 541 QUNGQPLVMVDGTLPELKPRLRAGRTLVIPVTMGFYVKNVNALACRYR 592  
 RESULT 3  
 AAU07424  
 ID AAU07424 standard; protein; 592 AA.  
 XX AAU07424;  
 AC  
 XX  
 XX 18-DEC-2001 (first entry)  
 DT  
 XX  
 XX Human heparanase-like protein splice variant #1.  
 DE  
 XX  
 XX Human: immunosuppressive; antiarthritic; antirheumatic; cytostatic;  
 KM antiproliferative; cardiact; vasotropic; cerebroprotective; nootropic;  
 KM neuroprotective; antibacterial; virucide; fungicide; ophthalmological;  
 KM extracellular matrix; ECM; autoimmune disease; rheumatoid arthritis;  
 KM hyperproliferative disorder; neoplasm; cardiovascular disorder;  
 KM cardiac arrest; cerebrovascular disorder; cerebral ischaemia; infection;  
 KM nervous system disorder; Alzheimer's disease; ocular disorder; sunburn;  
 KM wound healing; food additive; heparanase.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO200179253-A1.  
 PN  
 XX 25-OCT-2001.  
 PD  
 XX 11-APR-2001; 2001WO-US011643.  
 PF  
 XX 18-APR-2000; 2000US-0198123P.  
 PR  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX Fiscella M, Shi Y, Ebdner R, Ruben SM;  
 PI WPI, 2001-611720/70.  
 DR N-PDS; AAS13848.  
 XX  
 XX New nucleic acids encoding extracellular matrix polypeptides, for  
 PT diagnosing, treating, preventing or ameliorating human disorders and  
 PT disease, such as, autoimmune, hyperproliferative or cardiovascular  
 PT disorders.  
 PT  
 XX  
 XX Disclosure; Page 14; 308bp; English.  
 PS  
 XX  
 XX The invention relates to novel isolated polynucleotides (1) encoding  
 CC extracellular matrix (ECM) polypeptides. (1) and a polypeptide encoded by  
 CC (1) are used to prevent, treat or ameliorate a medical condition in e.g.  
 CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They  
 CC are also used in diagnosing a pathological condition or susceptibility to  
 CC a pathological condition. The antibodies to the polypeptides can also be  
 CC used in alleviating symptoms associated with the disorders and in  
 CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked  
 CC immunoassay assays (ELISA). Disorders which are diagnosed or treated  
 CC include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative  
 CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders  
 CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
 CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,  
 CC infections caused by bacteria, viruses and fungi and ocular disorders  
 CC e.g. corneal infection. The polypeptides can also be used to aid wound  
 CC healing and epithelial cell proliferation, to prevent skin aging due to

CC sunburn, to maintain organs before transplantation, for supporting cell  
 CC culture of primary tissues, to regenerate tissues and in chemotaxis. The  
 CC polypeptides can also be used as a food additive or preservative to  
 CC increase or decrease storage capabilities. The present sequence  
 CC represents the amino acid sequence of human heparanase-like protein,  
 CC splice variant #1  
 SQ Sequence 592 AA;  
 QY Query Match 99.7%; Score 3078; DB 4; Length 592;  
 Db Best Local Similarity 99.7%; Pred. No. 1,3e-303;  
 Matches 590; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MRVLCAPPEAMPSSNSRPPACIAPGALYIALLLHLSLSSQAGRRPVPVRAAGLKEKTL 60  
 Db 1 MRVLCAPPEAMPSSNSRPPACIAPGALYIALLLHLSLSSQAGRRPVPVRAAGLKEKTL 60  
 QY 61 ILDDVSTKNPVRTVNEFLSLQDPSIIHDGMLDFLSSKRLVTLARGLSPAFRFGGKRT 120  
 Db 61 ILDDVSTKNPVRTVNEFLSLQDPSIIHDGMLDFLSSKRLVTLARGLSPAFRFGGKRT 120  
 QY 121 DFLQFQNLRNPAKRGGRGPPYLYKNVEDIVSDVALDQKQCKLAQHPDWMLVLOREK 180  
 Db 121 DFLQFQNLRNPAKRGGRGPPYLYKNVEDIVSDVALDQKQCKLAQHPDWMLVLOREK 180  
 QY 181 AAQNHLYLKEQFNTYSNLLITARSIDLKYNFADCSGLHLIPALNALRPNNSNNSSS 240  
 Db 181 AAQNHLYLKEQFNTYSNLLITARSIDLKYNFADCSGLHLIPALNALRPNNSNNSSS 240  
 QY 241 ALSLKTSASGKKNISWELGNEPNNTTMRANVSGQLGKDYIQLKSLOPIRIYSRAS 300  
 Db 241 ALSLKTSASGKKNISWELGNEPNNTTMRANVSGQLGKDYIQLKSLOPIRIYSRAS 300  
 QY 301 LYGPNIQPRKQVYALDLDFPMKVGASTVDVAVTQHCYIDGRVAVKMDFLKRLDLTLDQ 360  
 Db 301 LYGPNIQPRKQVYALDLDFPMKVGASTVDVAVTQHCYIDGRVAVKMDFLKRLDLTLDQ 360  
 QY 361 IRKIQKVNTYTPGKIKWLEGVVTTSSAGTNNLSDSYIAGTFLMNTGMLANQSIDVIVR 420  
 Db 361 IRKIQKVNTYTPGKIKWLEGVVTTSSAGTNNLSDSYIAGTFLMNTGMLANQSIDVIVR 420  
 QY 421 HSPFDHGYNHLVDONENPLPDYWLSLYKRLIGPKYLAHVAGLQRPKPRPVIRDKLRI 480  
 Db 421 HSPFDHGYNHLVDONENPLPDYWLSLYKRLIGPKYLAHVAGLQRPKPRPVIRDKLRI 480  
 QY 481 YACCTNHNHNHYVRSITLFTIINLHRSRKKIKLAGTRDKLVHGYLLQPYGQEGLSKSV 540  
 Db 481 YACCTNHNHNHYVRSITLFTIINLHRSRKKIKLAGTRDKLVHGYLLQPYGQEGLSKSV 540  
 QY 541 QUNGQPLVMVDGTLPELKPRLRAGRTLVIPVTMGFYVKNVNALACRYR 592  
 Db 541 QUNGQPLVMVDGTLPELKPRLRAGRTLVIPVTMGFYVKNVNALACRYR 592  
 RESULT 4  
 AAB85215  
 ID AAB85215 standard; protein; 592 AA.  
 XX AAB85215;  
 AC  
 XX 07-SEP-2001 (first entry)  
 DT  
 XX  
 XX Heparanase-like protein Hpa2 splice variant #1.  
 DE  
 XX  
 XX Heparanase; splice variant; homologue; heparanase-like protein; Hpa2;  
 KM cytostatic; neuroprotective; cerebroprotective; immunosuppressive;  
 KM antiproliferative; nootropic; antiinflammatory; antiarthritic; antiasthmatic;  
 KM antidiabetic; antiarteriosclerotic; vulnerary.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 237

FT /label= unknown  
 FT /note="encoded by ANC"  
 XX  
 XX MO200146392-A2.  
 XX  
 XX 28-JUN-2001.  
 XX  
 XX 21-DEC-2000; 2000MO-GB004963.  
 XX  
 XX 22-DEC-1999; 99GB-00030392.  
 XX  
 XX 07-APR-2000; 2000GB-00008713.  
 XX  
 XX (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
 XX  
 XX Mckenzie EA, Stamps AC, Terrett JA, Tyson KL;  
 XX  
 XX WPI; 2001-418056/44.  
 XX  
 XX N-PSDB; AAM22671.  
 XX  
 XX Novel homologs of heparanase, present in three splice variants, useful  
 PT for identifying agents that modulate heparanase, useful in the treatment  
 PT and/or prophylaxis of abnormal levels of heparanase.  
 XX  
 XX Claim 1; Fig 1; 97pp; English.  
 PS  
 XX The invention provides a homologue to heparanase which is present in  
 CC three splice variants. The heparanase homologue polypeptide is useful in  
 CC the treatment of a human or non-human animal or for use in diagnosis.  
 CC Vectors comprising the heparanase homologue polynucleotides are useful in  
 CC the transformation or transfection of a prokaryotic or eukaryotic host.  
 CC The modulators of the polypeptide are useful in the manufacture of a  
 CC medicament for the treatment and/or prophylaxis of a condition/disease  
 CC associated with abnormal levels of the heparanase homologue, including  
 CC cancer, central nervous system (CNS) and neurodegenerative diseases,  
 CC cardiovascular diseases such as restenosis following angioplasty and  
 CC atherosclerosis, autoimmune diseases, psoriasis, lupus erythematosus,  
 CC allografts, inflammatory diseases, arthritis, vascular restenosis, tumour  
 CC growth and progression, asthma, Alzheimer's disease, diabetic  
 CC retinopathy, wound healing and inflammation. The polypeptide is also  
 CC useful in diagnosis and research. The present sequence represents the  
 CC amino acid sequence of the largest splice variant of the heparanase-like  
 CC protein Hpa2 of the invention  
 XX  
 XX Sequence 592 AA;  
 SQ  
 Query Match 99.4%; Score 3071; DB 4; Length 592;  
 Best Local Similarity 99.5%; Pred. No. 6.9e-303; Indels 0; Gaps 0;  
 Matches 589; Conservative 1; Mismatches 2;  
 QY 1 MRVLCAPELAMPSSNSRPPACLAGALYLLALHTLSLSSQAGDRRPLPVDRAAGLKEKTL 60  
 Db 1 MRVLCAPELAMPSSNSRPPACLAGALYLLALHTLSLSSQAGDRRPLPVDRAAGLKEKTL 60  
 QY 61 ILIDVSTKPNRTNENPLSLQDPSITIHGMDLDFSSKRLVLTARLSAPFRFGKRT 120  
 Db 61 ILIDVSTKPNRTNENPLSLQDPSITIHGMDLDFSSKRLVLTARLSAPFRFGKRT 120  
 QY 121 DFLQFQNLNRPFAKSGPSPDYLYKNYEDDIVRSVALDLQCKGCKIAQHBDVVLQREK 180  
 Db 121 DFLQFQNLNRPFAKSGPSPDYLYKNYEDDIVRSVALDLQCKGCKIAQHBDVVLQREK 180  
 QY 181 AAQNHVLLKQESNTTNSLILTRASDLKYNPDCGGLHIFPLNLRPNPNNSWSSS 240  
 Db 181 AAQNHVLLKQESNTTNSLILTRASDLKYNPDCGGLHIFPLNLRPNPNNSWSSS 240  
 QY 241 ALSLKTSASKKYVIMELGNEPNRYTHGRANVNSQLGQDYIQLSLQPIRIYSRAS 300  
 Db 241 ALSLKTSASKKYVIMELGNEPNRYTHGRANVNSQLGQDYIQLSLQPIRIYSRAS 300  
 QY 301 LYGNIGRPRKNVALLDGFNKVAGSTVDANTWGHCTIDGAVYKVMPLKTRLLDTSDQ 360  
 Db 301 LYGNIGRPRKNVALLDGFNKVAGSTVDANTWGHCTIDGAVYKVMPLKTRLLDTSDQ 360

QY 361 IRKIQKVNNTYTPGKKIWEAGVTTSAAGTNNISDSYAGFLMNTLGMANOGDIVIR 420  
 Db 361 IRKIQKVNNTYTPGKKIWEAGVTTSAAGTNNISDSYAGFLMNTLGMANOGDIVIR 420  
 QY 421 HSPFDHGYNHLVQONENPLPDYMLSLYKRLIDPKYLAHVAGLQRRPRGRYIRDKLR 480  
 Db 421 HSPFDHGYNHLVQONENPLPDYMLSLYKRLIDPKYLAHVAGLQRRPRGRYIRDKLR 480  
 QY 481 YAHCTNNHNNHYVAGSITLFIINLHRSRKKIKLAGTLRDLVHQYLLQPYGQGLSKSV 540  
 Db 481 YAHCTNNHNNHYVAGSITLFIINLHRSRKKIKLAGTLRDLVHQYLLQPYGQGLSKSV 540  
 QY 541 QINQGPVMDDDGTLPELKRPLRAGTLYIPPTWGFYVKNVNLACRYR 592  
 Db 541 QINQGPVMDDDGTLPELKRPLRAGTLYIPPTWGFYVKNVNLACRYR 592  
 RESULT 5  
 AAE18326  
 ID AAE18326 standard; protein; 582 AA.  
 XX  
 AC AAE18326;  
 XX  
 XX 07-MAY-2002 (first entry)  
 DT  
 XX Human heparanase-2AB splice variant protein.  
 DE  
 XX Human; heparanase-2AB; Hep-2; wound healing; angiogenesis; restenosis;  
 KW atherosclerosis; neurodegenerative disease; inflammation; prostate;  
 KW viral infection; autoimmune lesion; renal failure; pancreatic cancer;  
 KW dystrophic muscular disease; heart disease; gene therapy; enzyme.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200204645-A2.  
 PN  
 XX 17-JAN-2002.  
 PD  
 XX 12-JUL-2001; 2001WO-BP008094.  
 PF  
 XX 12-JUL-2000; 2000EP-00202442.  
 PR  
 XX (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.  
 PA  
 XX David G. Duerr J;  
 PI  
 XX WPI; 2002-171719/22.  
 DR  
 XX N-PSDB; AAD29202.  
 DR  
 XX Heparanase-2 polypeptides and polynucleotides, useful for useful in wound  
 PT healing, angiogenesis, and for treating restenosis, atherosclerosis,  
 PT inflammation, neurodegenerative diseases, and viral infections.  
 XX  
 PS Claim 1; Page 38-40; 54pp; English.  
 XX  
 XX The invention relates to human heparanase-2 (Hep-2) polypeptides and  
 CC polynucleotides. Heparanase-2 protein is useful in wound healing,  
 CC angiogenesis, restenosis, atherosclerosis, neurodegenerative diseases,  
 CC inflammation and viral infections, as well as in neutralising plasma  
 CC heparin as a potential replacement of protamine. Antiheparanase-2  
 CC antibodies may be used for immunodetection and diagnosis of  
 CC micrometastases, autoimmune lesions, renal failure in biopsy specimens,  
 CC plasma samples and body fluids. Molecules, which can agonise or  
 CC antagonise heparanase 2 catalytic activity may also be used as a  
 CC medicament. Polymorphisms in the polynucleotide sequence are useful in  
 CC the identification of individuals having a predisposition to acquire  
 CC diseases resulting from an increased or decreased expression of their  
 CC activity. Such molecules can be used to treat pancreatic cancer,  
 CC dystrophic muscular diseases and or heart diseases. Polynucleotides of  
 CC the invention are used in gene therapy. The present sequence is human  
 CC heparanase-2AB splice variant protein  
 XX  
 XX Sequence 582 AA;  
 SQ

Query Match 97.4%; Score 3007; DB 5; Length 582;  
 Best Local Similarity 99.1%; Pred. No. 2.2e-296;  
 Matches 577; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 11 MSSNSRPPACAPGALYALALHLHLSSQAGDRRPLPVDRAAGLKEKTLILDVSTNPF 70  
 1 MSSNSRPPACAPGALYALALHLHLSSQAGDRRPLPVDRAAGLKEKTLILDVSTNPF 60  
 DB 71 VRTVENFLSLQDPSIHDGMDFLSSKRLVTLARGSPAFIRFGKRTDFLOQNTN 130  
 61 VRTVENFLSLQDPSIHDGMDFLSSKRLVTLARGSPAFIRFGKRTDFLOQNTN 120  
 QY 131 PAKSRGPGPDYLYKAYEDDIVRSVVALPKQCKIAQHPDVMYLQREKAQOMLVLLK 190  
 121 PAKSRGPGPDYLYKAYEDDIVRSVVALPKQCKIAQHPDVMYLQREKAQOMLVLLK 180  
 QY 191 EOPSNTYSMLITARSGLKYNFADCSGLHLIPALNALRRNNNSWSSSALSILKYSAS 250  
 181 EOPSNTYSMLITARSGLKYNFADCSGLHLIPALNALRRNNNSWSSSALSILKYSAS 240  
 QY 251 KKNINISWELGEPNNRTWAGRAVNSQKQYIOCKSLQPIRYSRASLIGPNIGRPR 310  
 241 KKNINISWELGEPNNRTWAGRAVNSQKQYIOCKSLQPIRYSRASLIGPNIGRPR 300  
 QY 311 KXVIALLDGFMKAVGSTVDAVTWQHICYIDGRVVKVMDFLKTRLLDLSQIRKIQVNT 370  
 301 KXVIALLDGFMKAVGSTVDAVTWQHICYIDGRVVKVMDFLKTRLLDLSQIRKIQVNT 360  
 QY 371 YTFGKKIMLEGVTTTSAAGTNNISDSYAAAGFLMNTLGMLANOGIDVIRHSFPDHGYNH 430  
 361 YTFGKKIMLEGVTTTSAAGTNNISDSYAAAGFLMNTLGMLANOGIDVIRHSFPDHGYNH 420  
 QY 431 LVQNFNPDPYVLSLYKELIGPKYLAHVAGLQRPGRVIRDKLRIYAHCTNNHNH 490  
 421 LVQNFNPDPYVLSLYKELIGPKYLAHVAGLQRPGRVIRDKLRIYAHCTNNHNH 480  
 QY 491 NYVRSITITFINLHRSRKKIKLAGLRDLVHQYLLQPYGQGLKSKSVQANGPVMV 550  
 481 NYVRSITITFINLHRSRKKIKLAGLRDLVHQYLLQPYGQGLKSKSVQANGPVMV 540  
 QY 551 DDGTLBELKRPRLAGRTLVIPVTMGFFVVKVNNALACRYR 592  
 541 DDGTLBELKRPRLAGRTLVIPVTMGFFVVKVNNALACRYR 582  
 DB

RESULT 6  
 AA97633  
 ID AA97633 standard; protein; 538 AA.  
 AC AA97633;  
 DT 20-APR-2001 (first entry)  
 DE Human heparanase, hnhp1 p99 form, protein sequence.  
 XX Heparanase; hnhp1; wound healing; angiogenesis; restenosis; Scurvy;  
 KW atherosclerosis; inflammation; pulmonary disease; Alzheimer's disease;  
 KM neurodegenerative disease; Creutzfeldt-Jakob disease; viral infection;  
 XX gene therapy; human.  
 OS Homo sapiens.  
 XX  
 FT Key Location/Qualifiers  
 FT Misc-difference 305 /note="encoded by GAC"  
 FT  
 XX WO200100643-A2.  
 XX 04-JAN-2001.  
 XX 19-JUN-2000; 2000MO-IL000358.  
 XX

PR 25-JUN-1999; 99US-0140801P.  
 XX (INST-) INSIGHT STRATEGY & MARKETING LTD.  
 XX  
 XX Pecker I, Michael I, Itzhaki H;  
 XX  
 XX WPI; 2001-137930/14.  
 XX DR N-PSDB; AAA91098.  
 XX  
 PT New polynucleotides and polypeptides that are distantly homologous to  
 PT heparanase, useful in wound healing, as well as in gene therapy protocols  
 PT for angiogenesis, restenosis, atherosclerosis, or inflammation.  
 XX  
 PS Claim 10; Page 61-62; 67pp; English.  
 XX  
 CC This sequence represents a heparanase of the invention. The heparanase  
 CC DNA and protein sequences are useful in wound healing, angiogenesis,  
 CC restenosis, atherosclerosis, inflammation, pulmonary disease,  
 CC neurodegenerative diseases (such as Scurvy, Alzheimer's disease, and  
 CC Creutzfeldt-Jakob disease) or viral infections. The heparanase coding  
 CC sequence is particularly useful in gene therapy  
 XX

Sequence 538 AA:  
 QY 1 MRVLCAPPEAPSSNSRPPACAPGALYALALHLHLSSQAGDRRPLPVDRAAGLKEKTL 60  
 DB 1 MRVLCAPPEAPSSNSRPPACAPGALYALALHLHLSSQAGDRRPLPVDRAAGLKEKTL 60  
 QY 61 ILDVDSTKNPRTYVENFLSLQDPSIHDGMDFLSSKRLVTLARGSPAFIRFGKRT 120  
 DB 61 ILDVDSTKNPRTYVENFLSLQDPSIHDGMDFLSSKRLVTLARGSPAFIRFGKRT 120  
 QY 121 DFLQPNLRNPAKRGPGPDYLYKAYEDDIVRSVVALPKQCKIAQHPDVMYLQREK 180  
 DB 121 DFLQPNLRNPAKRGPGPDYLYKAYEDDIVRSVVALPKQCKIAQHPDVMYLQREK 180  
 QY 181 AAQMLVLLKEQFNTYSNLTITARSGLKYNFADCSGLHLIPALNALRRNNNSWSSS 240  
 DB 181 AAQMLVLLKEQFNTYSNLTITARSGLKYNFADCSGLHLIPALNALRRNNNSWSSS 240  
 QY 241 ALSILKYSASAKKNNISWELGEPNNRTWAGRAVNSQKQYIOCKSLQPIRYSRAS 300  
 DB 241 ALSILKYSASAKKNNISWELGEPNNRTWAGRAVNSQKQYIOCKSLQPIRYSRAS 300  
 QY 301 LYGPNIGRPRKXVIALLDGFMKAVGSTVDAVTWQHICYIDGRVVKVMDFLKTRLLDLSQ 360  
 DB 301 LYGPNIGRPRKXVIALLDGFMKAVGSTVDAVTWQHICYIDGRVVKVMDFLKTRLLDLSQ 360  
 QY 361 IRKIQVNTYTFGKKIMLEGVTTTSAAGTNNISDSYAAAGFLMNTLGMLANOGIDVIR 420  
 DB 361 IRKIQVNTYTFGKKIMLEGVTTTSAAGTNNISDSYAAAGFLMNTLGMLANOGIDVIR 420  
 QY 421 HSFPDHGYNHLDQNFNPDPYVLSLYKELIGPKYLAHVAGLQRPGRVIRDKLRI 480  
 DB 421 HSFPDHGYNHLDQNFNPDPYVLSLYKELIGPKYLAHVAGLQRPGRVIRDKLRI 480  
 QY 481 YAHCTNNHNHYVRSITITFINLHRSRKKIKLAGLRDLVHQYLLQPYGQGLKSKSV 540  
 DB 481 YAHCTNNHNHYVRSITITFINLHRSRKKIKLAGLRDLVHQYLLQPYGQGLKSKSV 540  
 QY 541 QANGQPLVMVDGTLBELKRPRLAGRTLVIPVTMGFFVVKVNNALACRYR 592  
 DB 541 QANGQPLVMVDGTLBELKRPRLAGRTLVIPVTMGFFVVKVNNALACRYR 538

RESULT 7  
 AA85216  
 ID AA85216 standard; protein; 534 AA.  
 XX

AC AAB85216;  
 XX 07-SEP-2001 (first entry)  
 DE Heparanase-like protein Hpa2 splice variant #2.  
 XX Heparanase; splice variant; homologue; heparanase-like protein; Hpa2;  
 KM cytosolic; neuroprotective; cerebroprotective; immunosuppressive;  
 KM antiproliferative; nototropic; antiinflammatory; antiarthritic; antiaesthetic;  
 KM antidiabetic; antiarteriosclerotic; vulnery.  
 OS Homo sapiens.  
 XX MO200146392-A2.  
 XX 28-JUN-2001.  
 PD 21-DEC-2000; 2000WO-GB004963.  
 PF 22-DEC-1999; 59GB-00030392.  
 PR 07-APR-2000; 2000GB-00008713.  
 XX (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
 PA McKenzie EA, Stamps AC, Terrett JA, Tyson XL;  
 XX MPI; 2001-418056/44.  
 DR N-PSDB; AAB22672.  
 XX Novel homologs of heparanase, present in three splice variants, useful  
 PT for identifying agents that modulate heparanase, useful in the treatment  
 PT and/or prophylaxis of abnormal levels of heparanase.  
 XX Claim 1; Fig 2; 97pb; English.  
 XX The invention provides a homologue to heparanase which is present in  
 CC three splice variants. The heparanase homologue polypeptide is useful in  
 CC the treatment of a human or non-human animal or for use in diagnosis.  
 CC Vectors comprising the heparanase homologue polynucleotides are useful in  
 CC the transformation or transfection of a prokaryotic or eukaryotic host.  
 CC The modulators of the polypeptide are useful in the manufacture of a  
 CC medicament for the treatment and/or prophylaxis of a condition/disease  
 CC associated with abnormal levels of the heparanase homologue, including  
 CC cancer, central nervous system (CNS) and neurodegenerative diseases,  
 CC cardiovascular diseases such as restenosis following angioplasty and  
 CC atherosclerosis, autoimmune diseases, psoriasis, lupus erythematosus,  
 CC allografts, inflammatory diseases, arthritis, vascular restenosis, tumour  
 CC growth and progression, asthma, Alzheimer's disease, diabetic  
 CC retinopathy, wound healing and inflammation. The polypeptide is also  
 CC useful in diagnosis and research. The present sequence represents the  
 CC amino acid sequence of the mid-sized splice variant of the heparanase-  
 CC like protein Hpa2 of the invention  
 XX Sequence 534 AA;  
 SO  
 Query March 88.6%; Score 2736; DB 4; Length 534;  
 Best Local Similarity 89.9%; Pred. No. 7.5e-269; Indels 58; Gaps 1;  
 Matches 533; Conservative 1; Mismatches 1;  
 QY 1 MRVLCAPPEAMPSSNSRPPACLAGALYLLALHLSTISSQAGRRPLVDRAAGLKEKTL 60  
 DB 1 MRVLCAPPEAMPSSNSRPPACLAGALYLLALHLSTISSQAGRRPLVDRAAGLKEKTL 60  
 QY 1 IILDVSTKFNVRATNEMFLSLQDPSITIHGMDFLSSKRLVTLAAGLSPAFRFGSKRT 120  
 DB 61 IILDVSTKFNVRATNEMFLSLQDPSITIHGMDFLSSKRLVTLAAGLSPAFRFGSKRT 120  
 QY 61 IILDVSTKFNVRATNEMFLSLQDPSITIHGMDFLSSKRLVTLAAGLSPAFRFGSKRT 120  
 DB 121 DFLQFQMLRNPAPKSGRGGPDYLLKNYEDDIVSDVALDKQKCKIAQHPDVLVLQREX 180  
 QY 121 DFLQFQMLRNPAPKSGRGGPDYLLKNYEDDIVSDVALDKQKCKIAQHPDVLVLQREX 180  
 DB 121 DFLQFQMLRNPAPKSGRGGPDYLLKNYEDDIVSDVALDKQKCKIAQHPDVLVLQREX 180  
 QY 181 AAGMHLVLKEQPSNTYSNLTLT----- 203  
 DB 181 AAGMHLVLKEQPSNTYSNLTLT----- 203

DB 181 AAGMHLVLKEQPSNTYSNLTLT----- 203  
 QY 241 ALSLLKYSASKKNYISWEIGNEPNRYRTMGRANVSGQKGYIOLKSLQPIRISRAS 300  
 DB 204 -----EPNNRYRTMGRANVSGQKGYIOLKSLQPIRISRAS 242  
 QY 301 LYGPNIGRPKKNVIALLDGFMKAGSTVDAYVMQHCYIDGRVVKWDFLXRLDLDLSQ 360  
 DB 243 LYGPNIGRPKKNVIALLDGFMKAGSTVDAYVMQHCYIDGRVVKWDFLXRLDLDLSQ 302  
 QY 361 IRKIQKVVNTYTPGKKIMLBEVTTTSAGSTNNLSDSYAGFLMNTLGMLANQGDIVVR 420  
 DB 303 IRKIQKVVNTYTPGKKIMLBEVTTTSAGSTNNLSDSYAGFLMNTLGMLANQGDIVVR 362  
 QY 421 HSPFDGYNHLVQNTNPLPDYVLSLYRRLIGPKYLAHVAGLQKPRPRGVIRDKLRI 480  
 DB 363 HSPFDGYNHLVQNTNPLPDYVLSLYRRLIGPKYLAHVAGLQKPRPRGVIRDKLRI 422  
 QY 481 YACCTNNHNNVYRGSTLFTIINLARSKKIKLAGTLRDKLVHGYLLQYGGGLKSKSV 540  
 DB 423 YACCTNNHNNVYRGSTLFTIINLARSKKIKLAGTLRDKLVHGYLLQYGGGLKSKSV 482  
 QY 541 QLNQGPLVWDDGTLPELKRPLRAGRTLVIPVTMGFYVVKVNALACRYR 592  
 DB 483 QLNQGPLVWDDGTLPELKRPLRAGRTLVIPVTMGFYVVKVNALACRYR 534  
 RESULT 8  
 ID ABE69310 standard; protein; 534 AA.  
 XX ABE69310;  
 AC ABE69310;  
 XX 20-JAN-2003 (first entry)  
 DE Human polypeptide SEQ ID NO 1357.  
 XX Human; genome mapping; gene therapy; food supplement; virus; fungus;  
 KM cell-proliferative disorder; neurodegenerative disease; bacterial;  
 KM Parkinson's disease; Alzheimer's disease; autoimmune disease;  
 KM multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;  
 KM arthritis; cytotoxic; immunomodulator; nototropic; neuroprotective;  
 KM antiparkinsonian; antidiabetic; immunosuppressive; dermatological;  
 KM haemostatic; vulnery; fungicide; antibacterial; virotoxic; protozoacide;  
 KM antiarthritic.  
 XX Homo sapiens.  
 XX WO200270539-A2.  
 XX 12-SEP-2002.  
 PD 05-MAR-2002; 2002WO-US005095.  
 PF 05-MAR-2001; 2001US-00799451.  
 XX (HYSE-) HYSEQ INC.  
 XX Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;  
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,  
 PI Wehman T, Wang J, Wang D, Drmanac RT;  
 XX MPI; 2002-759812/82.  
 DR N-PSDB; ABE11527.  
 XX New polynucleotides comprising sequences assembled from expressed  
 PT sequence tags (ESTs), useful for treating cell-proliferative,  
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet  
 PT or coagulation disorders.  
 PS Claim 9; SEQ ID NO 1357; 1012pb + Sequence Listing; English.  
 CC The invention relates to an isolated polynucleotide (I) comprising a

CC nucleotide sequence selected from any of 948 sequences (AB211119-  
 CC AB212066) or their mature protein coding portion, active domain coding  
 CC protein or complementary sequences. The polynucleotides are useful for  
 CC identifying expressed genes or for physical mapping of human genome. The  
 CC encoded polypeptides (ABP68902-ABP68949) are useful as molecular weight  
 CC markers, as a food supplement, for generating antibodies, in medical  
 CC imaging, screening and diagnostic assays and for treating cell-  
 CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's  
 CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,  
 CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders, liver  
 CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver  
 CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),  
 CC arthritis, etc. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 534 AA:  
 SQ  
 Query Match 88.6%; Score 2736; DB 5; Length 534;  
 Best Local Similarity 89.9%; Pred. No. 7.5e-269;  
 Matches 532; Conservative 1; Mismatches 1; Indels 58; Gaps 1;  
 QY 1 MRYVLCAPPEAMPSSNSRPPACLPALYALTLHLTSLSSQAGRRLPDRAGLKEKTL 60  
 Db 1 MRYVLCAPPEAMPSSNSRPPACLPALYALTLHLTSLSSQAGRRLPDRAGLKEKTL 60  
 QY 61 ILIDVSTKNTVNTVNEFLSLQDPSTIHDGMLDFSSKRLVTLARGLSPAFRFSGKRT 120  
 Db 61 ILIDVSTKNTVNTVNEFLSLQDPSTIHDGMLDFSSKRLVTLARGLSPAFRFSGKRT 120  
 QY 121 DFLQFQNLNRPASRGSGPDDYIKNYEDDIIVASDVALDKQCKIAQHPDVLVLOREK 180  
 Db 121 DFLQFQNLNRPASRGSGPDDYIKNYEDDIIVASDVALDKQCKIAQHPDVLVLOREK 180  
 QY 181 AAOQHVLTLKEQFSNTNTSNLITPARSLDKYFPADSGHLIPALNALRNPNSWSSS 240  
 Db 181 AAOQHVLTLKEQFSNTNTSNLITPARSLDKYFPADSGHLIPALNALRNPNSWSSS 240  
 QY 241 ALSLIKYSASKKNIISWELNEPNNVETMGRAVNSQLGKDYIOLKSLIPIRISRAS 300  
 Db 241 ALSLIKYSASKKNIISWELNEPNNVETMGRAVNSQLGKDYIOLKSLIPIRISRAS 300  
 QY 204 -----EPNNVETMGRAVNSQLGKDYIOLKSLIPIRISRAS 242  
 Db 204 -----EPNNVETMGRAVNSQLGKDYIOLKSLIPIRISRAS 242  
 QY 301 LYGNIRPRKRVATALLDGMKVAGSTVDVWMOHCIDRIVYKVMDFELRLDLTSDQ 360  
 Db 301 LYGNIRPRKRVATALLDGMKVAGSTVDVWMOHCIDRIVYKVMDFELRLDLTSDQ 360  
 QY 243 LYGNIRPRKRVATALLDGMKVAGSTVDVWMOHCIDRIVYKVMDFELRLDLTSDQ 302  
 Db 243 LYGNIRPRKRVATALLDGMKVAGSTVDVWMOHCIDRIVYKVMDFELRLDLTSDQ 302  
 QY 361 IRKIQKVVNTYTPGKTIWLEGVVTTSAGTNNLSDSYAAQFLMNTLGMLANOGIDVIR 420  
 Db 361 IRKIQKVVNTYTPGKTIWLEGVVTTSAGTNNLSDSYAAQFLMNTLGMLANOGIDVIR 420  
 QY 421 HSFEDGYNHLYVDGNFPLDPYMLSLYKRLIGPKYLAHVAGIQRPRGRVIRDLRI 480  
 Db 421 HSFEDGYNHLYVDGNFPLDPYMLSLYKRLIGPKYLAHVAGIQRPRGRVIRDLRI 480  
 QY 363 HSFEDGYNHLYVDGNFPLDPYMLSLYKRLIGPKYLAHVAGIQRPRGRVIRDLRI 422  
 Db 363 HSFEDGYNHLYVDGNFPLDPYMLSLYKRLIGPKYLAHVAGIQRPRGRVIRDLRI 422  
 QY 481 YAHCTNNHNNHYVAGSTLFTLINLHRSRKIKLAGTLRDLKHQYLQPYGOBGLSKSV 540  
 Db 481 YAHCTNNHNNHYVAGSTLFTLINLHRSRKIKLAGTLRDLKHQYLQPYGOBGLSKSV 540  
 QY 423 YAHCTNNHNNHYVAGSTLFTLINLHRSRKIKLAGTLRDLKHQYLQPYGOBGLSKSV 482  
 Db 423 YAHCTNNHNNHYVAGSTLFTLINLHRSRKIKLAGTLRDLKHQYLQPYGOBGLSKSV 482  
 QY 541 QLNQGPLVWDDGTLPELKPRPLAAGRTLVPTTGMFYVKNNAACRYR 592  
 Db 541 QLNQGPLVWDDGTLPELKPRPLAAGRTLVPTTGMFYVKNNAACRYR 592  
 QY 483 QLNQGPLVWDDGTLPELKPRPLAAGRTLVPTTGMFYVKNNAACRYR 534  
 Db 483 QLNQGPLVWDDGTLPELKPRPLAAGRTLVPTTGMFYVKNNAACRYR 534  
 RESULT 9  
 AAMS0337  
 ID AAMS0337 standard; protein: 534 AA.  
 XX AAMS0337;  
 AC AAMS0337;  
 XX  
 DT 04-FEB-2002 (first entry)  
 XX  
 DE Human prepro-heparanase II.  
 XX

KW Heparanase II; human; vulnerary; angiogenesis inhibitor;  
 KW antiinflammatory; cytostatic; therapy; diagnosis.  
 XX  
 XX Homo sapiens.  
 OS  
 XX Key  
 XX Peptide  
 FT 1. .41  
 FT /label= Signal\_peptide  
 FT 42. .534  
 FT /label= Mature protein  
 FT /note= "specifically claimed in Claim 23 (b) "  
 FT Protein  
 FT 42. .129  
 FT /label= 8\_kDa subunit  
 FT /note= "specifically claimed in Claim 23 (c) "  
 FT Modified-site  
 FT 66. .68  
 FT /note= "O-phosphorylated by protein kinase C"  
 FT Modified-site  
 FT 97. .99  
 FT /note= "O-phosphorylated by protein kinase C"  
 FT Modified-site  
 FT 98. .100  
 FT /note= "O-phosphorylated by protein kinase C"  
 FT Modified-site  
 FT 116  
 FT /note= "Amidated"  
 FT Protein  
 FT 162. .534  
 FT /label= 50\_kDa subunit  
 FT /note= "specifically claimed in Claim 23 (f) "  
 FT Modified-site  
 FT 217. .219  
 FT /note= "Asn is N-glycosylated"  
 FT Modified-site  
 FT 315  
 FT /note= "Amidated"  
 FT Modified-site  
 FT 334. .336  
 FT /note= "Asn is N-glycosylated"  
 FT Modified-site  
 FT 449. .451  
 FT /note= "O-phosphorylated by protein kinase C"  
 FT Modified-site  
 FT 458. .560  
 FT /note= "O-phosphorylated by protein kinase C"  
 PN WO200101569-A2.  
 XX  
 PD 01-NOV-2001.  
 XX  
 PF 17-APR-2001; 2001WO-US010804.  
 XX  
 PR 20-APR-2000; 2000US-0199072P.  
 XX  
 PA (PHAA ) PHARMACIA & UPJOHN CO.  
 XX  
 PI Heinrichson RL, Bienkowski MJ;  
 XX WPI; 2002-04102/05.  
 DR N-PSDB; AA170705.  
 XX  
 PT Novel heparanase II polypeptide useful for identifying agents with alter  
 PT heparanase activity and for accelerating wound healing, blocking  
 PT angiogenesis or inflammation.  
 PS Claim 23 (a); Fig 1; 65pp; English.  
 XX  
 XX The present sequence is that of novel human prepro-heparanase II, a  
 CC parologue of human heparanase I. The sequence was deduced from isolated  
 CC cDNA clones obtained by database screening (see AA170705) Heparanase II  
 CC is a secreted protein that shows 43% identity at the amino acid level to  
 CC heparanase I. The prepro-protein is processed to remove a 41-amino acid  
 CC leader peptide, and further processed to remove internal amino acids,  
 CC yielding the 8 kDa (amino acids 42-129) and 50 kDa (amino acid 162-534)  
 CC subunits of the heparanase II enzyme. Heparanase I and II have a non-  
 CC overlapping expression pattern in human tissues and each may serve tissue  
 CC specific functional roles. The invention provides heparanase II nucleic  
 CC acids, vectors, host cells, polypeptides and antibodies. Polypeptides  
 CC comprising amino acids 42-534, 42-129, 42-161, 130-534 and 162-534 of the  
 CC present sequence, and nucleic acids encoding them, are specifically  
 CC claimed. Heparanase II is useful for identifying an agent that alters  
 CC heparanase activity. Such as agent is used in a claimed method for  
 CC treating a disease state. Inhibiting heparanase II activity is useful for



CC treating or preventing metastasis, cancer, CNS and neurodegenerative  
 CC diseases, inflammation and cardiovascular disease such as restenosis  
 CC following angioplasty and atherosclerosis, Heparanase II is useful for  
 CC accelerating wound healing, blocking angiogenesis, degradation of heparin  
 CC and neutralization of heparin's anticoagulant properties during surgery.  
 CC Heparanase or an agent that enhances heparanase activity can also be  
 CC infused into the vasculature to block accumulation and diapedesis of  
 CC neutrophils at sites of inflammation with or without added domains to  
 CC confer selectivity in delivery

XX Sequence 534 AA:

Query Match 88.6%; Score 2736; DB 5; Length 534;  
 Best Local Similarity 89.9%; Pred. No. 7,5e-269;  
 Matches 532; Conservative 1; Mismatches 1; Indels 58; Gaps 1;

QY 1 MRVLCAPPEAMPSSNSRPACLAGALYALILHLISLSQAGRRPLPYDRAAGLKEKTL 60  
 DB 1 MRVLCAPPEAMPSSNSRPACLAGALYALILHLISLSQAGRRPLPYDRAAGLKEKTL 60  
 QY 61 ILIDVSTKPNFRTVNEFSLQDPSIIHDGMLDPLSKRLVTLARGLSPAFRFQGRKT 120  
 DB 61 ILIDVSTKPNFRTVNEFSLQDPSIIHDGMLDPLSKRLVTLARGLSPAFRFQGRKT 120  
 QY 121 DFIQFQNLRRPAPSRGPGPDYLLKNYEDIVASDVALKQKCKIAQHPDVLVLQREK 180  
 DB 121 DFIQFQNLRRPAPSRGPGPDYLLKNYEDIVASDVALKQKCKIAQHPDVLVLQREK 180  
 QY 181 AAGMHVLKQFQSNFNTSNLILFARSLDKLNPADCSGLILFALNALRRPNNSWSSS 240  
 DB 181 AAGMHVLKQFQSNFNTSNLILFARSLDKLNPADCSGLILFALNALRRPNNSWSSS 240  
 QY 241 ALSLLKYSASKYNNISWELGNEPNNTYRMHGRAVNSQIGKDYTLQKSLLOPIRYSRAS 300  
 DB 241 ALSLLKYSASKYNNISWELGNEPNNTYRMHGRAVNSQIGKDYTLQKSLLOPIRYSRAS 300  
 QY 301 LYGPNIGRPKNVIALIDGFMKVAAGSTVAVTWQHCYIDGRVYKWMDFLKTRLDLSQ 360  
 DB 301 LYGPNIGRPKNVIALIDGFMKVAAGSTVAVTWQHCYIDGRVYKWMDFLKTRLDLSQ 360  
 QY 361 IRKIQKVVNTYTPGKKIMLEGVVTTSGGTNNLSDSYAAGFLMLNTLGMANQSIDVIR 420  
 DB 361 IRKIQKVVNTYTPGKKIMLEGVVTTSGGTNNLSDSYAAGFLMLNTLGMANQSIDVIR 420  
 QY 421 HSFPHDGYNLVQONRPLPDYMLSLYKRLIGPKYLAHVAGLQKRPGRVIRDKLRI 480  
 DB 421 HSFPHDGYNLVQONRPLPDYMLSLYKRLIGPKYLAHVAGLQKRPGRVIRDKLRI 480  
 QY 481 YACHTNHNANNVYRGSITLFTIIMLHRSRKKIKLAGTLRDLVHQYLLQPYGQGLKSKSV 540  
 DB 481 YACHTNHNANNVYRGSITLFTIIMLHRSRKKIKLAGTLRDLVHQYLLQPYGQGLKSKSV 540  
 QY 541 QLNGQPLVMVDGTLPELKRPRLRAGTTLVTPYTMGFYVKNVNLACRIR 592  
 DB 541 QLNGQPLVMVDGTLPELKRPRLRAGTTLVTPYTMGFYVKNVNLACRIR 592  
 QY 483 QLNGQPLVMVDGTLPELKRPRLRAGTTLVTPYTMGFYVKNVNLACRIR 534  
 DB 483 QLNGQPLVMVDGTLPELKRPRLRAGTTLVTPYTMGFYVKNVNLACRIR 534

RESULT 10  
 AAE18327  
 ID AAE18327 standard; protein: 528 AA.

XX AAE18327;  
 AC AAE18327;  
 DT 07-MAY-2002 (first entry)

XX Human heparanase-2A splice variant protein.

XX Human, heparanase-2A; Hep-2; wound healing; angiogenesis; restenosis;  
 KW atherosclerosis; neurodegenerative disease; inflammation; procarcane;  
 KM viral infection; autoimmune lesion; renal failure; pancreatic cancer;  
 KM dystrophic muscular disease; heart disease; gene therapy; enzyme.

OS Homo sapiens.

XX WO200204645-A2.  
 XX 17-JAN-2002.  
 XX 12-JUL-2001; 2001WO-EP008094.  
 XX 12-JUL-2000; 2000EP-00202442.  
 XX (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

XX David G, Duert J;  
 XX WPI; 2002-171719/22.  
 XX N-PSDB; AAD29204.

PT Heparanase-2 polypeptides and polynucleotides, useful for useful in wound  
 PT healing, angiogenesis, and for treating restenosis, atherosclerosis,  
 PT inflammation, neurodegenerative diseases, and viral infections.

PS Disclosure: Page 45-46; 54pp; English.

CC The invention relates to human heparanase-2 (Hep-2) polypeptides and  
 CC polynucleotides. Heparanase-2 protein is useful in wound healing,  
 CC angiogenesis, restenosis, atherosclerosis, neurodegenerative diseases,  
 CC inflammation and viral infections, as well as in neutralizing plasma  
 CC heparin as a potential replacement of procaine. Antiheparanase-2  
 CC antibodies may be used for immunodetection and diagnosis of  
 CC micrometastases, autoimmune lesions, renal failure in biopsy specimens,  
 CC plasma samples and body fluids. Molecules, which can agonize or  
 CC antagonize heparanase 2 catalytic activity may also be used as a  
 CC medicament. Polymorphisms in the polynucleotide sequence are useful in  
 CC the identification of individuals having a predisposition to acquire  
 CC diseases resulting from a increased or decreased expression of their  
 CC activity. Such molecules can be used to treat pancreatic cancer,  
 CC dystrophic muscular diseases and or heart diseases. Polynucleotides of  
 CC the invention are used in gene therapy. The present sequence is human  
 CC heparanase-2A splice variant protein

XX Sequence 528 AA:

Query Match 88.1%; Score 2720; DB 5; Length 528;  
 Best Local Similarity 90.4%; Pred. No. 3.1e-267;  
 Matches 526; Conservative 1; Mismatches 1; Indels 54; Gaps 1;

QY 11 MPSSNSRPACLAGALYALILHLISLSQAGRRPLPYDRAAGLKEKTLILIDVSTKPN 70  
 DB 1 MPSSNSRPACLAGALYALILHLISLSQAGRRPLPYDRAAGLKEKTLILIDVSTKPN 60  
 QY 71 VRTNENFSLQDPSIIHDGMLDPLSKRLVTLARGLSPAFRFQGRKDPLOFQNLRN 130  
 DB 71 VRTNENFSLQDPSIIHDGMLDPLSKRLVTLARGLSPAFRFQGRKDPLOFQNLRN 120  
 QY 61 VRTNENFSLQDPSIIHDGMLDPLSKRLVTLARGLSPAFRFQGRKDPLOFQNLRN 120  
 DB 61 VRTNENFSLQDPSIIHDGMLDPLSKRLVTLARGLSPAFRFQGRKDPLOFQNLRN 120  
 QY 131 PAKSRGPGPDYLLKNYEDIVASDVALKQKCKIAQHPDVLVLQREKRAQMHVLVLK 190  
 DB 131 PAKSRGPGPDYLLKNYEDIVASDVALKQKCKIAQHPDVLVLQREKRAQMHVLVLK 190  
 QY 121 PAKSRGPGPDYLLKNYEDIVASDVALKQKCKIAQHPDVLVLQREKRAQMHVLVLK 139  
 DB 121 PAKSRGPGPDYLLKNYEDIVASDVALKQKCKIAQHPDVLVLQREKRAQMHVLVLK 139  
 QY 191 EGFNTYENLILFARSLDKLNPADCSGLILFALNALRRPNNSWSSSALSLKXAS 250  
 DB 191 EGFNTYENLILFARSLDKLNPADCSGLILFALNALRRPNNSWSSSALSLKXAS 186  
 QY 140 -----ARSIDKLYNFPADCSGLILFALNALRRPNNSWSSSALSLKXAS 186  
 DB 140 -----ARSIDKLYNFPADCSGLILFALNALRRPNNSWSSSALSLKXAS 186  
 QY 251 KKYNIWELGNEPNNTYRMHGRAVNSQIGKDYTLQKSLLOPIRYSRASLYGPNIGRPR 310  
 DB 251 KKYNIWELGNEPNNTYRMHGRAVNSQIGKDYTLQKSLLOPIRYSRASLYGPNIGRPR 310  
 QY 187 KKYNIWELGNEPNNTYRMHGRAVNSQIGKDYTLQKSLLOPIRYSRASLYGPNIGRPR 246  
 DB 187 KKYNIWELGNEPNNTYRMHGRAVNSQIGKDYTLQKSLLOPIRYSRASLYGPNIGRPR 246  
 QY 311 KKVIALIDGFMKVAAGSTVAVTWQHCYIDGRVYKWMDFLKTRLDLSQIDRIKIQKVVNT 370  
 DB 311 KKVIALIDGFMKVAAGSTVAVTWQHCYIDGRVYKWMDFLKTRLDLSQIDRIKIQKVVNT 306  
 QY 247 KKVIALIDGFMKVAAGSTVAVTWQHCYIDGRVYKWMDFLKTRLDLSQIDRIKIQKVVNT 306  
 DB 247 KKVIALIDGFMKVAAGSTVAVTWQHCYIDGRVYKWMDFLKTRLDLSQIDRIKIQKVVNT 306  
 QY 371 YTPGKKIMLEGVVTTSGGTNNLSDSYAAGFLMLNTLGMANQSIDVIRHSFPHDGYNH 430  
 DB 307 YTPGKKIMLEGVVTTSGGTNNLSDSYAAGFLMLNTLGMANQSIDVIRHSFPHDGYNH 366

QY 431 LVONFNPDPYMLSLYKRLIGPKYLAHVAGLQKRPGRVIRDPKLRITAHCTNHHN 490  
 DB 367 LVONFNPDPYMLSLYKRLIGPKYLAHVAGLQKRPGRVIRDPKLRITAHCTNHHN 426  
 QY 491 NYVRSITLFIINLHRSRKKIKLAGTLRDLVHQYLLQPYGQGLSKSVOLNGQPLVWV 550  
 DB 427 NYVRSITLFIINLHRSRKKIKLAGTLRDLVHQYLLQPYGQGLSKSVOLNGQPLVWV 486  
 QY 551 DDGTLPELKRPLRAGRTIYIPPTMGFPYVKNVNNALACRYR 592  
 DB 487 DDGTLPELKRPLRAGRTIYIPPTMGFPYVKNVNNALACRYR 528

RESULT 11  
 AAB84664 standard; protein; 492 AA.  
 ID AAB84664;  
 AC AAB84664;  
 XX  
 XX 05-SEP-2001 (first entry)  
 DE Amino acid sequence of human heparanase-like polypeptide.  
 XX  
 XX Human; heparanase-like polypeptide; gene therapy; cancer; angiogenesis;  
 KW trauma; autoimmune disease; skin disease; cardiovascular disease;  
 KW nervous system disease; inflammation; arthritis; genitalia;  
 KW male fertility; erectile dysfunction.  
 XX  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Misc-difference 407  
 FT note="unspecified residue encoded by KCA"  
 XX  
 PN MO200148161-A2.  
 XX  
 PD 05-JUL-2001.  
 XX  
 PF 18-DEC-2000; 2000MO-EP012909.  
 XX  
 PR 23-DEC-1999; 99EP-00125831.  
 XX  
 PA (SCHD) SCHERING AG.  
 XX  
 PT Siemeister G, Weiss B;  
 XX  
 DR WPI; 2001-418259/44.  
 DR N-PSDB; AAH28347.  
 XX  
 PT Human Heparanase-like polynucleotide encoding polypeptides useful for  
 PT modulating expression of the polypeptide and for treating cancer, cancer  
 PT metastasis, aberrant angiogenesis by gene therapy technique.  
 XX  
 PS Claim 9; Page 30; 30pp; English.  
 XX  
 XX The present sequence represents a human heparanase-like polypeptide.  
 CC Heparanase-like polynucleotides are useful as a source of probes, primers  
 CC and antisense molecules, and in gene therapy. Heparanase-like  
 CC polynucleotides and polypeptides are useful for treating several  
 CC disorders e.g., cancer, cancer metastasis. The oligonucleotides are also  
 CC useful as diagnostic markers for the diagnosis of disorder such as  
 CC cancer, cancer metastasis and aberrant angiogenesis. They may also act as  
 CC diagnostic markers for diagnosis of disorder such as cancer, cancer  
 CC metastasis and aberrant angiogenesis. The heparanase polypeptides and  
 CC polynucleotides are also useful for treating trauma, autoimmune diseases,  
 CC skin diseases, cardiovascular diseases, nervous system diseases, and  
 CC inflammation including arthritis. Since the polynucleotide is  
 CC preferentially expressed in male genitalia, modulation of its expression  
 CC and/or activity may be used for medical intervention in male genitalia  
 CC function that is male fertility control, erectile dysfunction  
 CC  
 XX Sequence 492 AA;

Query Match 81.7%; Score 2522; DB 4; Length 492;  
 Best Local Similarity 89.1%; Pred. No. 4e-247;  
 Matches 490; Conservative 0; Mismatches 2; Indels 58; Gaps 1;

QY 43 DRRPLVDRAAGLKEKTLILLDVSTKQPVRTVNFSLQIDPSIIDDGWLDFSSKRLV 102  
 DB 1 DRRPLVDRAAGLKEKTLILLDVSTKQPVRTVNFSLQIDPSIIDDGWLDFSSKRLV 60  
 QY 103 TLARGLSPAPLRFGGKRTDFLOFQNLARNPAKSRGPGPDYLLKMYEDDIASVADLKOK 162  
 DB 61 TLARGLSPAPLRFGGKRTDFLOFQNLARNPAKSRGPGPDYLLKMYEDDIASVADLKOK 120  
 QY 163 GCKIAQHPDVLVLQREKAOQHLVLLKEQPSNTYSNLIITARSLDKLYNFADCSGLHLI 222  
 DB 121 GCKIAQHPDVLVLQREKAOQHLVLLKEQPSNTYSNLIITARSLDKLYNFADCSGLHLI 161  
 QY 223 PALNALRNPNNSSNLSLKSASAKTYNISWELNEBNYRTMGRAVNGSOLGKD 282  
 DB 162 -----EPNNYRTMGRAVNGSOLGKD 182  
 QY 283 YIQLSLQPIRIYRASLYGPNIGRPKNYIALIDGFMKVASGVDAVTWQHCVIDGRV 342  
 DB 183 YIQLSLQPIRIYRASLYGPNIGRPKNYIALIDGFMKVASGVDAVTWQHCVIDGRV 242  
 QY 343 VKVMDFLKTRLDITSDQIRKIQKVNTYTPGKKIWEQVTTGAGTNNISDSYAGFL 402  
 DB 243 VKVMDFLKTRLDITSDQIRKIQKVNTYTPGKKIWEQVTTGAGTNNISDSYAGFL 302  
 QY 403 WNTLGMANOGIDVIRHSFDFGQYHLVONFNPDPYMLSLYKRLIGPKYLAHVAV 462  
 DB 303 WNTLGMANOGIDVIRHSFDFGQYHLVONFNPDPYMLSLYKRLIGPKYLAHVAV 362  
 QY 463 GLQKRPGRVIRDPKLRITAHCTNHHNNYVRSITLFIINLHRSRKKIKLAGTLRDLV 522  
 DB 363 GLQKRPGRVIRDPKLRITAHCTNHHNNYVRSITLFIINLHRSRKKIKLAGTLRDLV 422  
 QY 523 HOYLLQPYGQGLSKSVOLNGQPLVWDDGTLPELKRPLRAGRTIYIPPTMGFPYVW 582  
 DB 423 HOYLLQPYGQGLSKSVOLNGQPLVWDDGTLPELKRPLRAGRTIYIPPTMGFPYVW 482  
 QY 583 NVNALACRYR 592  
 DB 483 NVNALACRYR 492

RESULT 12  
 AAY97634 standard; protein; 480 AA.  
 ID AAY97634;  
 AC AAY97634;  
 XX  
 XX 20-APR-2001 (first entry)  
 DE Human heparanase, hnp1 pns form, protein sequence.  
 XX  
 XX Heparanase; hnp1; wound healing; angiogenesis; restenosis; Scrape;  
 KW atherosclerosis; inflammation; pulmonary disease; Alzheimer's disease;  
 KW neurodegenerative disease; Creutzfeldt-Jakob disease; viral infection;  
 KW gene therapy; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO200100643-A2.  
 XX  
 PD 04-JAN-2001.  
 XX  
 PF 19-JUN-2000; 2000MO-IL000358.  
 XX  
 PR 25-JUN-1999; 99US-0140801P.  
 XX  
 PA (INST-) INSIGHT STRATEGY & MARKETING LTD.  
 XX  
 PI Pecker I, Michal I, Itzhaki H;

```

XX  WPI: 2001-137930/14.
DR  N-PSDB; AAA91099.
XX
XX  New polynucleotides and polypeptides that are distantly homologous to
PT  heparanase, useful in wound healing, as well as in gene therapy protocols
PT  for angiogenesis, restenosis, atherosclerosis, or inflammation.
XX
XX  Claim 10; Page 63; 67pp; English.
XX
XX  This sequence represents a heparanase of the invention. The heparanase
CC  DNA and protein sequences are useful in wound healing, angiogenesis,
CC  restenosis, atherosclerosis, inflammation, pulmonary diseases,
CC  neurodegenerative diseases (such as scrapie, Alzheimer's disease, and
CC  Creutzfeldt-Jakob disease) or viral infections. The heparanase coding
CC  sequence is particularly useful in gene therapy
XX
XX  Sequence 480 AA:
SQ
Query Match      79.2%; Score 2446; DB 4; Length 480;
Best Local Similarity 80.9%; Pred. No. 2.1e-239;
Matches 479; Conservative 1; Mismatches 0; Indels 112; Gaps 1;

OY  1 MRVLCAPFPEAMPSSNSRPPACIAPGALYALILHLISSSQAGDRRPLVDRAAGLKEKTL 60
DB  1 MRVLCAPFPEAMPSSNSRPPACIAPGALYALILHLISSSQAGDRRPLVDRAAGLKEKTL 60
OY  61 ILDDSTKPNVPTVENFLSLQDPSIHDGMDLFLSKRLVTLARGSPAFIRFGKRT 120
DB  61 ILDDSTKPNVPTVENFLSLQDPSIHDGMDLFLSKRLVTLARGSPAFIRFGKRT 120
OY  121 DFLQONTLRNPAKSRGGPGDYLLKNYEDDIVRSVALDKOKGCKIAQHPDVMVLQREK 180
DB  121 DFLQONTLRNPAKSRGGPGDYLLKNYEDDIVRSVALDKOKGCKIAQHPDVMVLQREK 180
OY  121 DFLQONTLRNPAKSRGGPGDYLLKNYEDDIVRSVALDKOKGCKIAQHPDVMVLQREK 180
DB  121 DFLQONTLRNPAKSRGGPGDYLLKNYEDDIVRSVALDKOKGCKIAQHPDVMVLQREK 180
OY  181 AAGMVLVLKEQFSNTYENLILTRASLDKLYNPAQSGMLIPALNLRNPNNSNNS 240
DB  181 AAGMVLVLKEQFSNTYENLILTRASLDKLYNPAQSGMLIPALNLRNPNNSNNS 240
OY  150 ----- 149
DB  150 ----- 149
OY  241 ALSLIKASAKKYNISWELGNEPNNTYRTMGRVAVNSQLGQYVQLKSLQPIRTYRAS 300
DB  150 ----- EPNNYRTMGRVAVNSQLGQYVQLKSLQPIRTYRAS 188
OY  301 LYGNIGRPKRVNIALDGMKVAAGSTYDAVTWQHCYDGVVYVMPFLKTRILDTLSDQ 360
DB  189 LYGNIGRPKRVNIALDGMKVAAGSTYDAVTWQHCYDGVVYVMPFLKTRILDTLSDQ 248
OY  361 IRKIQKVVNTYTPGKKIMLEGVTTTSACTNNISDSYAAGFLMNTLGMLANQGIIVVIR 420
DB  249 IRKIQKVVNTYTPGKKIMLEGVTTTSACTNNISDSYAAGFLMNTLGMLANQGIIVVIR 308
OY  421 HSPFHGNNHVDONFNLDPYVLSLYKRLIGRVLAHVHAGIQRKPRPRVIRDLKRI 480
DB  309 HSPFHGNNHVDONFNLDPYVLSLYKRLIGRVLAHVHAGIQRKPRPRVIRDLKRI 368
OY  481 YAHCTNHNHNVVRSITLFTINLHRSRKKIKLAGTLDKLVHGYLLQPYQGEGLKSKSV 540
DB  369 YAHCTNHNHNVVRSITLFTINLHRSRKKIKLAGTLDKLVHGYLLQPYQGEGLKSKSV 428
OY  541 QLNQGPLVMVDDGTPELKPRPLPAGRTLVIPVTMGVYVYVNNALACRRR 592
DB  429 QLNQGPLVMVDDGTPELKPRPLPAGRTLVIPVTMGFVYVXVNNALACRRR 480

RESULT 13
AAU07418
ID  AAU07418 standard; protein; 480 AA.
XX
XX  AAU07418;
AC
XX  18-DEC-2001 (first entry)
DT
XX
XX  Novel human extracellular matrix (ECM) protein #1.
DE

```

```

XX  Human: immunosuppressive; antiarthritic; antipneumatic; cytosolic;
XX  antiproliferative; cardiac; vasotropic; cerebroprotective; nootropic;
XX  neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
XX  extracellular matrix; ECM; autoimmune disease; rheumatoid arthritis;
XX  hyperproliferative disorder; neoplasm; cardiovascular disorder;
XX  cardiac arrest; cerebrovascular disorder; cerebral ischemia; infection;
XX  nervous system disorder; Alzheimer's disease; ocular disorder; sunburn;
XX  wound healing; food additive.
XX
XX  Homo sapiens.
XX
XX  WO200179253-A1.
XX
XX  25-OCT-2001.
XX
XX  11-APR-2001; 2001WO-US011643.
XX
XX  18-APR-2000; 2000US-0198123P.
XX
XX  (HUMA-) HUMAN GENOME SCI INC.
XX
XX  Fiselcia M, Shi Y, Ebner R, Ruben SM;
XX  WPI: 2001-611720/70.
XX  N-PSDB; AAS13843.
XX
XX  New nucleic acids encoding extracellular matrix polypeptides, for
PT  diagnosing, treating, preventing or ameliorating human disorders and
PT  disease, such as, autoimmune, hyperproliferative or cardiovascular
PT  disorders.
XX
XX  Claim 1; Page 292-293; 308pp; English.
XX
XX  The invention relates to novel isolated polynucleotides (I) encoding
CC  extracellular matrix (ECM) polypeptides. (I) and a polypeptide encoded by
CC  (I) are used to prevent, treat or ameliorate a medical condition in e.g.
CC  humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They
CC  are also used in diagnosing a pathological condition or susceptibility to
CC  a pathological condition. The antibodies to the polypeptides can also be
CC  used in alleviating symptoms associated with the disorders and in
CC  diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
CC  immunosorbent assays (ELISA). Disorders which are diagnosed or treated
CC  include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC  disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC  e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischemia,
CC  angiogenesis, nervous system disorders e.g. Alzheimer's disease,
CC  infections caused by bacteria, viruses and fungi and ocular disorders
CC  e.g. corneal infection. The polypeptides can also be used to aid wound
CC  healing and epithelial cell proliferation, to prevent skin aging due to
CC  sunburn, to maintain organs before transplantation, for supporting cell
CC  culture of primary tissues, to regenerate tissues and in chemotaxis. The
CC  polypeptides can also be used as a food additive or preservative to
CC  increase or decrease storage capabilities. The present sequence
CC  represents the amino acid sequence of novel human extracellular matrix
CC  (ECM) protein #1
XX
XX  Sequence 480 AA:
SQ
Query Match      79.2%; Score 2446; DB 4; Length 480;
Best Local Similarity 80.9%; Pred. No. 2.1e-239;
Matches 479; Conservative 1; Mismatches 0; Indels 112; Gaps 1;

OY  1 MRVLCAPFPEAMPSSNSRPPACIAPGALYALILHLISSSQAGDRRPLVDRAAGLKEKTL 60
DB  1 MRVLCAPFPEAMPSSNSRPPACIAPGALYALILHLISSSQAGDRRPLVDRAAGLKEKTL 60
OY  61 ILDDSTKPNVPTVENFLSLQDPSIHDGMDLFLSKRLVTLARGSPAFIRFGKRT 120
DB  61 ILDDSTKPNVPTVENFLSLQDPSIHDGMDLFLSKRLVTLARGSPAFIRFGKRT 120
OY  121 DFLQONTLRNPAKSRGGPGDYLLKNYEDDIVRSVALDKOKGCKIAQHPDVMVLQREK 180
DB  121 DFLQONTLRNPAKSRGGPGDYLLKNYEDDIVRSVALDKOKGCKIAQHPDVMVLQREK 180

```

```

Db      121 DFLQFQNLRLNPAKSGRGGPGPYLYKAYED----- 149
Qy      181 AAQMHVLLKEQFSNTYSNLIITARSIDKLYNFADCSGLHLIPALNALRRPNNSWNSSS 240
Db      150 ----- 149
Qy      241 ALSLTKYSASKKYNISWELGNEPNRYTMHGRAVNSQLGKDYIQLKSLLOPIRISRAS 300
Db      150 -----EPNNYRTMHGRAVNSQLGKDYIQLKSLLOPIRISRAS 188
Qy      301 LYGNIGRPRKNVIALLDGFMKVASITVDAYTWQHCYIDGRVYKVMFLKTRLLDITLSDQ 360
Db      189 LYGNIGRPRKNVIALLDGFMKVASITVDAYTWQHCYIDGRVYKVMFLKTRLLDITLSDQ 248
Qy      361 IRKIQKVVNTYTPGKKIMLEGVVTTSAAGTNNLSDSYAAGFLMNTLGMLANOGIDVYIR 420
Db      249 IRKIQKVVNTYTPGKKIMLEGVVTTSAAGTNNLSDSYAAGFLMNTLGMLANOGIDVYIR 308
Qy      421 HSFEDHGVNHLVDQNFNPLDPYWLSLYKRLIGPKVLAVHAGLQKRPGRVIRDKLRI 480
Db      309 HSFEDHGVNHLVDQNFNPLDPYWLSLYKRLIGPKVLAVHAGLQKRPGRVIRDKLRI 368
Qy      481 YAHCTNHNHNHYVRSITLFIINLHRSRKKIKLAGTLDKLVHGYLLQPYGEGLSKSV 540
Db      369 YAHCTNHNHNHYVRSITLFIINLHRSRKKIKLAGTLDKLVHGYLLQPYGEGLSKSV 428
Qy      541 QLNQGPLVWVDGTLPELKPRLRAGRTLVIPVTMGFVYVYVNVNALACRYR 592
Db      429 QLNQGPLVWVDGTLPELKPRLRAGRTLVIPVTMGFVYVYVNVNALACRYR 480

RESULT 14
AAB85217
ID      AAB85217 standard; protein; 480 AA.
XX
XX      AAB85217;
AC
XX      07-SEP-2001 (first entry)
DT
XX
XX      Heparanase-like protein Hpa2 splice variant #3.
DE
XX
XX      Heparanase; splice variant; homologue; heparanase-like protein; Hpa2;
KM      cytosolic; neuroprotective; cerebroprotective; immunosuppressive;
KM      antiportatic; nootropic; antiinflammatory; antidiabetic; antiaschmatic;
XX      antidiabetic; antiarteriosclerotic; vulnerary.
OS
XX      Homo sapiens.
XX
XX      WO200146392-A2.
PN
XX
XX      28-JUN-2001.
PD
XX
XX      21-DEC-2000; 2000WO-GB004963.
PF
XX
XX      22-DEC-1999; 99GB-00030392.
PR      07-APR-2000; 2000GB-00008713.
XX
XX
XX      (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
PA
XX      McKenzie EA, Stamps AC, Terrett JA, Tyson KL,
PI
XX      WPI; 2001-418056/44.
DR      N-PSDB; AAH22673.
XX
XX
XX      Novel homologs of heparanase, present in three splice variants, useful
XX      for identifying agents that modulate heparanase, useful in the treatment
XX      and/or prophylaxis of abnormal levels of heparanase.
XX      Claim 1; Fig 3; 97p; English.
XX
XX      The invention provides a homologue to heparanase which is present in
XX      three splice variants. The heparanase homologue polypeptide is useful in
XX      the treatment of a human or non-human animal or for use in diagnosis.

```

```

CC      Vectors comprising the heparanase homologue polynucleotides are useful in
CC      the transformation or transfection of a prokaryotic or eukaryotic host.
CC      The modulators of the polypeptide are useful in the manufacture of a
CC      medicament for the treatment and/or prophylaxis of a condition/disease
CC      associated with abnormal levels of the heparanase homologue, including
CC      cancer, central nervous system (CNS) and neurodegenerative diseases,
CC      cardiovascular diseases such as restenosis following angioplasty and
CC      atherosclerosis, autoimmune diseases, psoriasis, lupus erythematosus,
CC      allografts, inflammatory diseases, arthritis, vascular restenosis, tumour
CC      growth and progression, asthma, Alzheimer's disease, diabetic
CC      retinopathy, wound healing and inflammation. The polypeptide is also
CC      useful in diagnosis and research. The present sequence represents the
CC      amino acid sequence of the smallest splice variant of the heparanase-
CC      like protein Hpa2 of the invention
XX
XX      Sequence 480 AA:
SQ
XX
XX      Query Match          79.2%; Score 2446; DB 4; Length 480;
XX      Best Local Similarity 80.9%; Pred. No. 2.1e-239;
XX      Matches 479; Conservative 1; Mismatches 0; Indels 112; Gaps 1;
Qy      1 MEVLCAFPKAPMSSNSRPPACIAPGALYIALIHLISLSSQAGDRPLPVDRAAGLKEKTL 60
Db      1 MEVLCAFPKAPMSSNSRPPACIAPGALYIALIHLISLSSQAGDRPLPVDRAAGLKEKTL 60
Qy      61 ILADVSTKNPVATVNEFLSLQIDPSIIHDGMDPLSSSKRLVTLARGLSPAFLRPGKRT 120
Db      61 ILDDVSTKNPVATVNEFLSLQIDPSIIHDGMDPLSSSKRLVTLARGLSPAFLRPGKRT 120
Qy      121 DFLQFQNLRLNPAKSGRGGPGPYLYKAYED----- 149
Db      121 DFLQFQNLRLNPAKSGRGGPGPYLYKAYED----- 149
Qy      181 AAQMHVLLKEQFSNTYSNLIITARSIDKLYNFADCSGLHLIPALNALRRPNNSWNSSS 240
Db      150 ----- 149
Qy      241 ALSLTKYSASKKYNISWELGNEPNRYTMHGRAVNSQLGKDYIQLKSLLOPIRISRAS 300
Db      150 -----EPNNYRTMHGRAVNSQLGKDYIQLKSLLOPIRISRAS 188
Qy      301 LYGNIGRPRKNVIALLDGFMKVASITVDAYTWQHCYIDGRVYKVMFLKTRLLDITLSDQ 360
Db      189 LYGNIGRPRKNVIALLDGFMKVASITVDAYTWQHCYIDGRVYKVMFLKTRLLDITLSDQ 248
Qy      361 IRKIQKVVNTYTPGKKIMLEGVVTTSAAGTNNLSDSYAAGFLMNTLGMLANOGIDVYIR 420
Db      249 IRKIQKVVNTYTPGKKIMLEGVVTTSAAGTNNLSDSYAAGFLMNTLGMLANOGIDVYIR 308
Qy      421 HSFEDHGVNHLVDQNFNPLDPYWLSLYKRLIGPKVLAVHAGLQKRPGRVIRDKLRI 480
Db      309 HSFEDHGVNHLVDQNFNPLDPYWLSLYKRLIGPKVLAVHAGLQKRPGRVIRDKLRI 368
Qy      481 YAHCTNHNHNHYVRSITLFIINLHRSRKKIKLAGTLDKLVHGYLLQPYGEGLSKSV 540
Db      369 YAHCTNHNHNHYVRSITLFIINLHRSRKKIKLAGTLDKLVHGYLLQPYGEGLSKSV 428
Qy      541 QLNQGPLVWVDGTLPELKPRLRAGRTLVIPVTMGFVYVYVNVNALACRYR 592
Db      429 QLNQGPLVWVDGTLPELKPRLRAGRTLVIPVTMGFVYVYVNVNALACRYR 480

RESULT 15
AAB18328
ID      AAB18328 standard; protein; 470 AA.
XX
XX      AAB18328;
AC
XX      07-MAY-2002 (first entry)
DT
XX
XX      Human heparanase-2B splice variant protein.
XX
XX      Human; heparanase-2B; Hep-2; wound healing; angiogenesis; restenosis;

```

KM atherosclerosis; neurodegenerative disease; inflammation; procaine;  
 KM viral infection; autoimmune lesion; renal failure; pancreatic cancer;  
 KM dystrophic muscular disease; heart disease; gene therapy; enzyme.  
 XX Homo sapiens.  
 XX MO200204645-A2.  
 XX 17-JAN-2002.  
 PD 12-JUL-2001; 2001MO-EP008094.  
 PF 12-JUL-2000; 2000EP-00202442.  
 PR 12-JUL-2000; 2000EP-00202442.  
 XX (VLA-) VLAMIS INTERUNIVERSITAIR INST BIOTECHNOG.  
 PA David G, Duerr J;  
 PI WPI: 2002-171719/22.  
 DR N-PSDB; MAD29205.  
 XX Heparanase-2 polypeptides and polymuclectides, useful for useful in wound  
 PT healing, angiogenesis, and for treating restenosis, atherosclerosis,  
 PT inflammation, neurodegenerative diseases, and viral infections.  
 PS Disclosure; Page 49-50; 54pp; English.  
 XX The invention relates to human heparanase-2 (Hep-2) polypeptides and  
 CC polymuclectides. Heparanase-2 protein is useful in wound healing,  
 CC angiogenesis, restenosis, atherosclerosis, neurodegenerative diseases,  
 CC inflammation and viral infections, as well as in neutralising plasma  
 CC heparin as a potential replacement of protamine. Antiheparanase-2  
 CC antibodies may be used for immunodetection and diagnosis of  
 CC micrometastases, autoimmune lesions, renal failure in biopsy specimens,  
 CC plasma samples and body fluids. Molecules, which can agonise or  
 CC antagonise heparanase 2 catalytic activity may also be used as a  
 CC medicament. Polymorphisms in the polymuclectide sequence are useful in  
 CC the identification of individuals having a predisposition to acquire  
 CC diseases resulting from a increased or decreased expression of their  
 CC activity. Such molecules can be used to treat pancreatic cancer,  
 CC dystrophic muscular diseases and or heart diseases. Polymuclectides of  
 CC the invention are used in gene therapy. The present sequence is human  
 CC heparanase-2B splice variant protein  
 CC  
 CC  
 SQ Sequence 470 AA;  
 Query Match 77.3%; Score 2388; DB 5; Length 470;  
 Best Local Similarity 80.4%; Pred. No. 1.6e-233;  
 Matches 468; Conservative 1; Mismatches 1; Indels 112; Gaps 1;  
 QY 11 MPSSNSRPPACLAGALYALILHLSTLSQAQDRRLPVDRAAGLKEKTLILLDVSTKXP 70  
 DB 1 MPSSNSRPPACLAGALYALILHLSTLSQAQDRRLPVDRAAGLKEKTLILLDVSTKXP 60  
 QY 71 VRTNENPFLSLQPLPSIIHGWMLDFLSKRLVTLARGLSPAPLRFGGKRTDFLOFQNLRY 130  
 DB 61 VRTNENPFLSLQPLPSIIHGWMLDFLSKRLVTLARGLSPAPLRFGGKRTDFLOFQNLRY 120  
 QY 131 PAKSRGGPGPDYLLKNYEDDIVRSVALDKQKCKIAQHPDVMVLVQREKRAQWHLVTLK 190  
 DB 121 PAKSRGGPGPDYLLKNYEDDIVRSVALDKQKCKIAQHPDVMVLVQREKRAQWHLVTLK 139  
 QY 191 EQFENTYSNLILTRASLDKLYNPAFCGSLHIFALNLRPNPNKSNWSSALSILTKYSAS 250  
 DB 140 ----- 139  
 QY 251 KKYNISWELGNEPNNYRTMGRAVNSQLGKDYIQLKSLQPIRIYSRSLYGPNIQRPR 310  
 DB 140 -----EPNNYRTMGRAVNSQLGKDYIQLKSLQPIRIYSRSLYGPNIQRPR 188  
 QY 311 KNVIALLDGFKKVGASTDAATWQHGYIDGRVYVVMDFLKTRLDLTSDQIRKIQRVNT 370  
 DB 189 KNVIALLDGFKKVGASTDAATWQHGYIDGRVYVVMDFLKTRLDLTSDQIRKIQRVNT 248

QY 371 YTPGKKIWLEGVVTTSAAGTNNLSDSYAAGFLMINTIGMLANQIDVIRHSFEDHGYNH 430  
 DB 249 YTPGKKIWLEGVVTTSAAGTNNLSDSYAAGFLMINTIGMLANQIDVIRHSFEDHGYNH 308  
 QY 431 LYDQNFNLPDYMTSLYKRLIGPKVLAVHAGIQRCPRRGRVIRDKLRIYAHCTNNHNH 490  
 DB 309 LYDQNFNLPDYMTSLYKRLIGPKVLAVHAGIQRCPRRGRVIRDKLRIYAHCTNNHNH 368  
 QY 491 NYVRGSITLFIINLRSHKIKLAGTLRDKLVHQYLLQPYGOEGLKSKSVQNLNGQPLVMV 550  
 DB 369 NYVRGSITLFIINLRSHKIKLAGTLRDKLVHQYLLQPYGOEGLKSKSVQNLNGQPLVMV 428  
 QY 551 DDGTLPELKPRLPRAGRVLVIPPVTMGFTYVKNVNALACRYR 592  
 DB 429 DDGTLPELKPRLPRAGRVLVIPPVTMGFTYVKNVNALACRYR 470

Search completed: May 6, 2004, 13:44:12  
 Job time : 63 secs

GenCore Version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2004, 13:43:17 ; Search time 23 Seconds  
(without alignments)  
1328.808 Million cell updates/sec

Title: US-10-088-676-2

Perfect score: 3088  
Sequence: 1 MRYLCAPPEAMPSSNSRPPA.....PTMGFYVKNVVALACRYR 592

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/2/1aa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PCUTS COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1165	37.7	536	US-09-930-218-2	Sequence 2, Appl1
2	1163.5	37.7	535	US-09-435-739-44	Sequence 44, Appl1
3	1163.5	37.7	535	US-09-930-218-1	Sequence 1, Appl1
4	1148.5	37.2	543	US-08-922-1708-10	Sequence 10, Appl1
5	1148.5	37.2	543	US-09-071-7398-2	Sequence 2, Appl1
6	1148.5	37.2	543	US-09-260-0388-2	Sequence 2, Appl1
7	1148.5	37.2	543	US-09-635-923-2	Sequence 2, Appl1
8	1148.5	37.2	543	US-09-487-716A-2	Sequence 2, Appl1
9	1148.5	37.2	543	US-09-322-977-2	Sequence 2, Appl1
10	1148.5	37.2	543	US-09-186-200-1	Sequence 1, Appl1
11	1148.5	37.2	543	US-09-435-739-10	Sequence 10, Appl1
12	1148.5	37.2	543	US-09-930-218-3	Sequence 3, Appl1
13	1148.5	37.2	543	US-09-435-739-14	Sequence 14, Appl1
14	1144.5	37.1	588	US-09-181-336-13	Sequence 13, Appl1
15	1144.5	37.1	588	US-09-601-777-2	Sequence 2, Appl1
16	1141.5	37.0	527	US-09-930-218-16	Sequence 16, Appl1
17	1141.5	37.0	530	US-09-252-586-2	Sequence 2, Appl1
18	1104.5	35.8	532	US-09-181-336-15	Sequence 15, Appl1
19	1033.5	33.5	523	US-09-930-218-4	Sequence 4, Appl1
20	984.5	31.9	380	US-09-181-336-19	Sequence 19, Appl1
21	963.5	31.2	380	US-09-181-336-17	Sequence 17, Appl1
22	1144	4.7	38	US-09-252-586-26	Sequence 26, Appl1
23	114.5	3.7	399	US-08-282-197C-54	Sequence 54, Appl1
24	114.5	3.7	837	US-09-350-234-14	Sequence 14, Appl1
25	114.5	3.3	837	US-09-603-311-14	Sequence 14, Appl1
26	100.5	3.3	1183	US-09-532-310B-6	Sequence 6, Appl1
27	97.5	3.2	774	US-09-328-352-5361	Sequence 5361, Ap

28	96.5	3.1	1874	US-09-331-403-2	Sequence 2, Appl1
29	96	3.1	990	US-08-392-625-20	Sequence 20, Appl1
30	96	3.1	990	US-08-466-961A-20	Sequence 20, Appl1
31	95.5	3.1	709	US-09-668-673B-3	Sequence 3, Appl1
32	95.5	3.1	849	US-08-405-392-18	Sequence 18, Appl1
33	95.5	3.1	849	US-08-487-691-18	Sequence 18, Appl1
34	95.5	3.1	849	US-08-189-738A-19	Sequence 19, Appl1
35	95.5	3.1	866	US-08-405-392-17	Sequence 17, Appl1
36	95.5	3.1	866	US-08-487-691-17	Sequence 17, Appl1
37	95.5	3.1	893	US-08-189-738A-18	Sequence 18, Appl1
38	95.5	3.1	905	US-08-405-392-2	Sequence 2, Appl1
39	95.5	3.1	905	US-08-487-691-2	Sequence 2, Appl1
40	95.5	3.1	905	US-08-666-221B-4	Sequence 4, Appl1
41	95.5	3.1	905	US-08-666-221B-10	Sequence 10, Appl1
42	95.5	3.1	905	US-08-189-738A-2	Sequence 2, Appl1
43	95	3.1	371	US-09-758-008-1	Sequence 1, Appl1
44	94	3.0	1103	US-08-455-543A-53	Sequence 53, Appl1
45	94	3.0	1103	US-08-223-305C-53	Sequence 53, Appl1

#### ALIGNMENTS

```

RESULT 1
US-09-930-218-2
Sequence 2, Application US/09930218
Patent No. 6677137
GENERAL INFORMATION:
APPLICANT: Goldschmidt, orit
APPLICANT: pecker, iris
APPLICANT: vlodavsky, israel
APPLICANT: israel, michael
TITLE OF INVENTION: AVIAN AND REPTILE DERIVED POLYNUCLEOTIDE ENCODING A POLYPEPTIDE H
FILE REFERENCE: 01/22335
CURRENT APPLICATION NUMBER: US/09/930, 218
PRIOR FILING DATE: 2001-08-16
PRIOR APPLICATION NUMBER: 09/666,390
PRIORITY FILING DATE: 2000-09-20
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 536
TYPE: PRT
ORGANISM: Rattus rattus
US-09-930-218-2

Query Match      37.7%, Score 1165, DB 4, Length 536,
Best Local Similarity 45.1%, Pred. No. 1e-111,
Matches 248, Conservative 87, Mismatches 187, Indels 28, Gaps 7;

QY 42 GRRRLPVDRAAG-LKEKTLILDVSTKNPVRTNENFLSLQDPSIHD-GMDLFTSSK 99
DB 12 GRLRLTGTPGTAAPYDQVVDLEFYTKLRFQSVSPSLITDASLATDRFLTLGSP 71
QY 100 RLVTLARGLSPAFLFSGKRTDPLQFONLRNPAYSRGGPGDYLYKAYEDDIVSDVALD 159
DB 72 RLRLARGLSPAFLFSGKRTDPLQFONLRNPAYSRGGPGDYLYKAYEDDIVSDVALD 159
QY 160 KQKGGKIQHPDVMVYLQREKAQMHTVLYLKEQSNSTNSMLILARSLDKYFNADCSGL 219
DB 124 -----RVSDAVLKLQWMPFO-ELHLLREQYQREFKNSYTSRSYDMLYFAKSRLL 175
QY 220 HLIPALNLRNPNNSWSSALSLTKYSASKYNIISWELGNEPNNTYRTMGRAVNSQL 279
DB 176 DIFELNLLRTPDRLMRSSNAQLLNCSKGYNISWELGNEPNNSFWKKAQISIDGL 235
QY 280 GSDYVQLSLQPIRIRYASLNGNTGRPKNTYALDGMKYVAGSVVDVVTMOHCYID 339
DB 236 GSDYVQLSLQPIRIRYASLNGNTGRPKNTYALDGMKYVAGSVVDVVTMOHCYID 339
QY 340 GGVVVMDFLKLRLDITLSDQIRKIQKVVNTYTPGKTIWLGVTTSAGTNTLSDSYAA 399
DB 340 GGVVVMDFLKLRLDITLSDQIRKIQKVVNTYTPGKTIWLGVTTSAGTNTLSDSYAA 399

```



```

Db 295 GRVATKEDFLSDVDLDTFLISVOKIKVTKEMTPGKKWLGETSAYGGAPLSTFPA 354
Qy 400 GFLMLNTLGLMANGIDVVRHSFPHGVNHLVDONFNPJPDYMLSLYKRLIGPVLV 459
Db 355 GFWMIDGLGSAQGLIEVVRQVFGAGNHLVDENFEDLPDGLSLFKLGGPVLS 414
Qy 460 HVAGLQRPKPPGVINDKRLIYAHCTNNHNNYVRSITLFIINLRSRKKI KLAGTLD 519
Db 415 RVKGPD-----RSKLRVYLCITVYHPRYREGDLTYVNLNVTGHLKLPFPMFS 465
Qy 520 KLHVHVLQPYGQEGKSKSVOLNGOPLVWDDGTLPEKRPRLRAGRTLVIPVPMGY 579
Db 466 RPVDKRLKFPSSDGLSKSVOLNGQTLKKNVDEQTLPALTEKLPAGSSLSVAFSYGFF 525
Qy 580 VKNVNAALAC 589
Db 526 VIRNAKIAAC 535

```

## RESULT 2

```

US-09-435-739-44
; Sequence 44, Application US/09435739
; Patent No. 6664105
; GENERAL INFORMATION:
; APPLICANT: Pecker, Iris
; APPLICANT: Vlodavsky, Israel
; APPLICANT: Feinstein, Elena
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY
; FILE REFERENCE: 00/20454
; CURRENT APPLICATION NUMBER: US/09/435, 739
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 44
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-435-739-44

```

```

Query Match 37.7%; Score 1163.5; DB 4; Length 535;
Best Local Similarity 46.2%; Pred. No. 1.4e-111;
Matches 244; Conservative 80; Mismatches 177; Indels 27; Gaps 6;

Qy 63 LDVSTKNPRTVNEFSLQDPSIHD-GWLDPLSSKRLVTLARGLSPAFLEFGKRTD 121
Db 33 LEFTYKRLRSLVSPFSLITIDAGLATDPRFLTLGSPRLALRGLSPAYLRFGGTKTD 92
Qy 122 FLQPNLRNPAKSGGGPDDYILKNYEDDIYRSVDALDKQKCKIAQHPDVMVLQREKA 181
Db 93 FLIF---DPKEPTSEERSYKWSQVNHDCRSEPV-----SAVLRKLOVEMP 137
Qy 182 AQMHLVLKEQPSNTYSLILITARSLDKLYNFADCSGLHLIFALNALRRPNNSWSSSA 241
Db 138 FQ-ELLRLREGYQKEFNSTYSRSVDMLYSFACSGDLIFGNMLLRTPDLRMSSNA 196
Qy 242 LSLKYSASKKNISWELGNEPNRYRTMGRVAVNGSLQGVYQLKSLLOPIRIYSASL 301
Db 197 QLLDLYCSSKGNISWELGNEPNRPMKKAHLIDGLQGEFVFLHKLQR-SAFQNAKL 255
Qy 302 YGNIGRPRKRVIALDGFMKVAGSTVDVAVTMOHCYIDGRVVKVMDLKTRLDPTLSDQI 361
Db 256 YGPRIGQPRKGVTLRLSFLKAGGEVIDSLTMHHYVINGRIATKEDFLSSDALDPTLSV 315
Qy 362 RKIQVAVNTYTPGKKIWEQVVTTSAGCTNNLSDSYAGFLMNTLGLMANGIDVYIRH 421
Db 316 QKILKVTKEITPPGKKVWLGETSAYGGAPLSTFPAAGFWMLDKGLSAQMGIEVVRQ 375
Qy 422 SFPDGYNHLVDONFNPJPDYMLSLYKRLIGPVLVVHAGLQRPKPPGVINDKRLIY 481
Db 376 VFFGAGNHLVDENFEDLPDGLSLFKLGGPVLSRVKGPD-----RSKLRVY 426
Qy 482 AHCNNHNNYVRSITLFIINLRSRKKIKLAGTLDKLVHGYLLOPYGQEGKSKSVQ 541

```

```

Db 427 LHCITVYHPRYQEGDLTYVNLNVTGHLKVPPELPRKKVDYTLKPSGPDGLSKSVQ 486
Qy 542 LNGQPLVWDDGTLPEKRPRLRAGRTLVIPVPMGYVKNVNAALAC 589
Db 487 LNGQTLKKNVDEQTLPALTEKLPAGSSLSVAFSYGFVIRNAKIAAC 534

```

## RESULT 3

```

US-09-930-218-1
; Sequence 1, Application US/09930218
; Patent No. 6677137
; GENERAL INFORMATION:
; APPLICANT: goldsmith, orit
; APPLICANT: pecker, iris
; APPLICANT: vlodavsky, israel
; APPLICANT: israel, michael
; TITLE OF INVENTION: AVIAN AND REPTILE DERIVED POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY
; FILE REFERENCE: 01/22335
; CURRENT APPLICATION NUMBER: US/09/930, 218
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 09/666,390
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-930-218-1

```

```

Query Match 37.7%; Score 1163.5; DB 4; Length 535;
Best Local Similarity 46.2%; Pred. No. 1.4e-111;
Matches 244; Conservative 80; Mismatches 177; Indels 27; Gaps 6;

Qy 63 LDVSTKNPRTVNEFSLQDPSIHD-GWLDPLSSKRLVTLARGLSPAFLEFGKRTD 121
Db 33 LEFTYKRLRSLVSPFSLITIDAGLATDPRFLTLGSPRLALRGLSPAYLRFGGTKTD 92
Qy 122 FLQPNLRNPAKSGGGPDDYILKNYEDDIYRSVDALDKQKCKIAQHPDVMVLQREKA 181
Db 93 FLIF---DPKEPTSEERSYKWSQVNHDCRSEPV-----SAVLRKLOVEMP 137
Qy 182 AQMHLVLKEQPSNTYSLILITARSLDKLYNFADCSGLHLIFALNALRRPNNSWSSSA 241
Db 138 FQ-ELLRLREGYQKEFNSTYSRSVDMLYSFACSGDLIFGNMLLRTPDLRMSSNA 196
Qy 242 LSLKYSASKKNISWELGNEPNRYRTMGRVAVNGSLQGVYQLKSLLOPIRIYSASL 301
Db 197 QLLDLYCSSKGNISWELGNEPNRPMKKAHLIDGLQGEFVFLHKLQR-SAFQNAKL 255
Qy 302 YGNIGRPRKRVIALDGFMKVAGSTVDVAVTMOHCYIDGRVVKVMDLKTRLDPTLSDQI 361
Db 256 YGPRIGQPRKGVTLRLSFLKAGGEVIDSLTMHHYVINGRIATKEDFLSSDALDPTLSV 315
Qy 362 RKIQVAVNTYTPGKKIWEQVVTTSAGCTNNLSDSYAGFLMNTLGLMANGIDVYIRH 421
Db 316 QKILKVTKEITPPGKKVWLGETSAYGGAPLSTFPAAGFWMLDKGLSAQMGIEVVRQ 375
Qy 422 SFPDGYNHLVDONFNPJPDYMLSLYKRLIGPVLVVHAGLQRPKPPGVINDKRLIY 481
Db 376 VFFGAGNHLVDENFEDLPDGLSLFKLGGPVLSRVKGPD-----RSKLRVY 426
Qy 482 AHCNNHNNYVRSITLFIINLRSRKKIKLAGTLDKLVHGYLLOPYGQEGKSKSVQ 541
Db 487 LNGQTLKKNVDEQTLPALTEKLPAGSSLSVAFSYGFVIRNAKIAAC 534

```

RESULT 4  
US-08-922-170B-10  
Sequence 10, Application US/08922170B  
Patent No. 5968822  
GENERAL INFORMATION:  
APPLICANT: Iris Pecker, Israel Vlodavsky and Elena  
Applicant: Feinstein  
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE  
TITLE OF INVENTION: HAVING HEPARANASE ACTIVITY AND EXPRESSION OF  
TITLE OF INVENTION: SAME IN TRANSFUSED CELLS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Mark M. Friedman c/o Robert Sheindein  
STREET: 2940 Birchtree lane  
CITY: Silver Spring  
STATE: Maryland  
COUNTRY: United States of America  
ZIP: 20906  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
COMPUTER: Twinhead\* Slimnote-890TX  
OPERATING SYSTEM: MS DOS version 6.2,  
OPERATING SYSTEM: Windows version 3.11  
SOFTWARE: Word for Windows version 2.0 converted to  
SOFTWARE: an ASCII file  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/922,170B  
FILING DATE: 2 SEP 1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Friedman, Mark M.  
REGISTRATION NUMBER: 33,883  
REFERENCE/DOCKET NUMBER: 910/1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 972-3-5625553  
TELEFAX: 972-3-5625554  
TELEX:  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 543  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-922-170B-10  
Query Match 37.2%; Score 1148.5; DB 2; Length 543;  
Best Local Similarity 43.4%; Pred. No. 5.3e-110;  
Matches 249; Conservative 82; Mismatches 190; Indels 53; Gaps 9;  
DB 18 PPAALAPALYALALHLHLSLSQAGDRRPLVDRAAGLKEKTLILLDVSTKQNVRTVNN 77  
20 PLGLPSGL-----PPA-----QAQVVDLDFFOEPLHVSFS 55  
DB 78 FLISQDPSLIHD-GWIDFLSSKRLVTLARGSLPAPLRGGKXTDPLQFQNLNPAKSRG 136  
56 FLSTVTDANATPRLIFLIGSPPLRLARGSLPAPLRGGKXTDPLIF---DPKKEST 111  
DB 137 GPGDYYLYKNVBDIVSDVALDKQKCKIAQ-HPDVMTVLOREKAQOMELVLLKQPSN 195  
112 FEERSVWQSVNQDI-----CKYGSIPDVEEKLRLIEWPYQEGT-LIREHYOK 158  
DB 196 TYSVLIITARSIDKLYNPAFCGSHLIPALNALRRNNNNSSSALSLKYSASKYNI 255  
159 KFKKSTYSRSSVDVLYTFANCSGDLIFGLNALRLRTADLQWNSNNAQLLDYCSSKGYNI 218  
DB 256 SWEIAGNPNRYRTMGRAVNGSOLGKDYIOLKSLIOPIRIYSRASIYGNIGRPKRVIA 315  
219 SWEIAGNPNRFKADLFINGSOIGEDYIQHKLIRK-STFKAKALXGPGVCGPRRTAK 277  
DB 316 LLDSFMKVASGTVDATVWQHGYIDGRVVKWMDFLKTRLLDTLSQIKRIQKVNTYTPGK 375

RESULT 5  
US-09-071-739B-2  
Sequence 2, Application US/09071739B  
Patent No. 617545  
GENERAL INFORMATION:  
APPLICANT: Iris Pecker et al.  
TITLE OF INVENTION: HEPARANASE SPECIFIC MOLECULAR PROBES  
TITLE OF INVENTION: AND THEIR USE IN RESEARCH AND MEDICAL  
TITLE OF INVENTION: APPLICATIONS  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Mark M. Friedman c/o Anthony Caetorina  
STREET: 20001 Jefferson Davis Highway, Suite 207  
CITY: Arlington  
STATE: Virginia  
COUNTRY: United States of America  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
COMPUTER: Twinhead\* Slimnote-890TX  
OPERATING SYSTEM: MS DOS version 6.2,  
OPERATING SYSTEM: Windows version 3.11  
SOFTWARE: Word for Windows version 2.0 converted to  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071,739B  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/922,180  
FILING DATE: September 2, 1997  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Friedman, Mark M.  
REGISTRATION NUMBER: 33,883  
REFERENCE/DOCKET NUMBER: 910/5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 972-3-5625553  
TELEFAX: 972-3-5625554  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 543  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-071-739B-2  
Query Match 37.2%; Score 1148.5; DB 3; Length 543;  
Best Local Similarity 43.4%; Pred. No. 5.3e-110;  
Matches 249; Conservative 82; Mismatches 190; Indels 53; Gaps 9;  
DB 18 PPAALAPALYALALHLHLSLSQAGDRRPLVDRAAGLKEKTLILLDVSTKQNVRTVNN 77

```

Db      20  FLGPIPSGAL-----PRPA-----QADVVDDLPFTOEPLHVS 55
Qy      78  FLSLODPSIIHD-GWIDFLSKRLVTLARGSPAPLRFGGKRTDPLQFONLRNPAKRG 136
Db      56  FLSTVIDANLMTDPRFILLGSPKRLTARGLSPAYLRFGGTKTDPLIF-----DPKEEST 111
Qy      137  GPGPDYLYKATVEDDLYRSDVALDKQCKCKIAQ-HPDYMVLQREKKAQMHVYLKQFON 195
Db      112  FEERSYQSQVNQD-----CKYGSIPDVEEKLRLWPQEQ-LRREHYOK 158
Qy      196  TYSNLIITARSIDLKLYNFADCSGLHLIFALNALRPNPNNSWSSSALSLKYSASKKYN 255
Db      159  KFKNSTYRSRSDVLYTFANCSGDLIFGLNALRTADLQWNSNACLLDYCSSKGYNI 218
Qy      256  SWEIGNEPNRTTHGRAVNGSQLDKDYIOLKSLQPIRISPSASLYGPNIGPRKQVIA 315
Db      219  SWEIGNEPNRSFLKKADIFINGSQIGEDYIOLHKLRLK-STFKNAKLYGPDVGQPRKRTAK 277
Qy      316  LLDGFMKAVGSTVDATVWQHCYIDGRVVKWMDFLKTRLLDPLSDQIRKIQKVNVTYTPGK 375
Db      278  MLKSFLLKAGGEVIDSVTHHYTLNKRTRATEDPLNPDVLDIFISSVQKVFQVESTREPK 337
Qy      376  KIMLEGVTTTSAAGTNNISDSYAGFLMNTLGMLANQIDVIRHSFDPHGYNHLVDON 435
Db      338  KVMIGETSSAYGGGAPILSDTFPAAGFWMLDKLGASAMGIEVWVRQVFFGAGNYHLVDEN 397
Qy      436  FNPILPDYMLSLYKRLIGPKYLAHVAGLQKRPGRVIRDKLRIYACHTNNHNNHYRG 495
Db      398  FDBLPDYMLSLFLKVLGKYLKVAASVQSKR-----KARVYLACTNTDNPYKRG 448
Qy      496  SITLFTINLHRSRKIKLAGTLDKLVHOYLIQPYGQGLSKYSQVQNGOPLWMDGTL 555
Db      449  DLTLYAINLHNTYKXLRFPYPSNKQYDKYLRLPGHGLSKSVQVQNGTLKWDQTL 508
Qy      556  PELKRPRLAAGRTIVIPVTMGVYVKNVNALAC 589
Db      509  PLMEKPLRPGSSLGILPASFYSFVIRNAKVAAC 542

```

```

RESULT 6
US-09-260-038B-2
; Sequence 2, Application US/09260038B
; Patent No. 6348344
; GENERAL INFORMATION:
; APPLICANT: Maly Ayal-Herskovitz et al.
; TITLE OF INVENTION: GENETICALLY MODIFIED CELLS AND METHODS FOR
; EXPRESSING RECOMBINANT HEPARANASE
; AND METHODS OF PURIFYING SAME
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mark M. Friedman c/o Anthony Castorina
; STREET: 2001 Jefferson Davis Highway, Suite 207
; CITY: Arlington
; STATE: Virginia
; COUNTRY: United States of America
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
; COMPUTER: Twinhead+ Slimnote-890TX
; OPERATING SYSTEM: MS DOS version 6.2,
; Windows version 3.11
; SOFTWARE: Word for Windows version 2.0 converted to
; an ASCII file
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/260,038B
; FILING DATE: 02-Mar-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/071,618
; FILING DATE: May 1, 1998
; APPLICATION NUMBER: 09/071,739
; FILING DATE: May 1, 1998

```

```

; APPLICATION NUMBER: 08/922,180
; FILING DATE: September 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Friedman, Mark M.
; REGISTRATION NUMBER: 33,883
; REFERENCE/DOCKET NUMBER: 910/16
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 972-3-562553
; TELEFAX: 972-3-562554
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 543
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

```

Query Match      37.2%; Score 1148.5; DB 4; Length 543;
Best Local Similarity 43.4%; Pred. No. 5.3e-110;
Matches 249; Conservative 82; Mismatches 190; Indels 53; Gaps 9;

```

```

Qy      18  PPACIAPGALVIALLLHLSLSSQAGDRRLPVDAAGLKEKTKLLDVTSTKNIPVRYTEN 77
Db      20  FLGPIPSGAL-----PRPA-----QADVVDDLPFTOEPLHVS 55
Qy      78  FLSLODPSIIHD-GWIDFLSKRLVTLARGSPAPLRFGGKRTDPLQFONLRNPAKRG 136
Db      56  FLSTVIDANLMTDPRFILLGSPKRLTARGLSPAYLRFGGTKTDPLIF-----DPKEEST 111
Qy      137  GPGPDYLYKATVEDDLYRSDVALDKQCKCKIAQ-HPDYMVLQREKKAQMHVYLKQFON 195
Db      112  FEERSYQSQVNQD-----CKYGSIPDVEEKLRLWPQEQ-LRREHYOK 158
Qy      196  TYSNLIITARSIDLKLYNFADCSGLHLIFALNALRPNPNNSWSSSALSLKYSASKKYN 255
Db      159  KFKNSTYRSRSDVLYTFANCSGDLIFGLNALRTADLQWNSNACLLDYCSSKGYNI 218
Qy      256  SWEIGNEPNRTTHGRAVNGSQLDKDYIOLKSLQPIRISPSASLYGPNIGPRKQVIA 315
Db      219  SWEIGNEPNRSFLKKADIFINGSQIGEDYIOLHKLRLK-STFKNAKLYGPDVGQPRKRTAK 277
Qy      316  LLDGFMKAVGSTVDATVWQHCYIDGRVVKWMDFLKTRLLDPLSDQIRKIQKVNVTYTPGK 375
Db      278  MLKSFLLKAGGEVIDSVTHHYTLNKRTRATEDPLNPDVLDIFISSVQKVFQVESTREPK 337
Qy      376  KIMLEGVTTTSAAGTNNISDSYAGFLMNTLGMLANQIDVIRHSFDPHGYNHLVDON 435
Db      338  KVMIGETSSAYGGGAPILSDTFPAAGFWMLDKLGASAMGIEVWVRQVFFGAGNYHLVDEN 397
Qy      436  FNPILPDYMLSLYKRLIGPKYLAHVAGLQKRPGRVIRDKLRIYACHTNNHNNHYRG 495
Db      398  FDBLPDYMLSLFLKVLGKYLKVAASVQSKR-----KARVYLACTNTDNPYKRG 448
Qy      496  SITLFTINLHRSRKIKLAGTLDKLVHOYLIQPYGQGLSKYSQVQNGOPLWMDGTL 555
Db      449  DLTLYAINLHNTYKXLRFPYPSNKQYDKYLRLPGHGLSKSVQVQNGTLKWDQTL 508
Qy      556  PELKRPRLAAGRTIVIPVTMGVYVKNVNALAC 589
Db      509  PLMEKPLRPGSSLGILPASFYSFVIRNAKVAAC 542

```

```

RESULT 7
US-09-635-923-2
; Sequence 2, Application US/09635923
; Patent No. 6426209
; GENERAL INFORMATION:
; APPLICANT: Maly Ayal-Herskovitz et al.
; TITLE OF INVENTION: GENETICALLY MODIFIED CELLS AND METHODS FOR
; EXPRESSING RECOMBINANT HEPARANASE
; AND METHODS OF PURIFYING SAME

```

NUMBER OF SEQUENCES: 25.  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Mark M. Friedman c/o Anthony Castorina  
 STREET: 2001 Jefferson Davis Highway, Suite 207  
 CITY: Arlington  
 STATE: Virginia  
 COUNTRY: United States of America  
 ZIP: 22202  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
 COMPUTER: PC  
 OPERATING SYSTEM: MS DOS version 6.2,  
 Windows version 3.11  
 SOFTWARE: Word for Windows version 2.0 converted to an ASCII file  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/635,923  
 FILING DATE: 10-Aug-2000  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/09/487,716  
 FILING DATE: 19-Jan-2000  
 APPLICATION NUMBER: 09/071,618  
 FILING DATE: May 1, 1998  
 APPLICATION NUMBER: 09/071,739  
 FILING DATE: May 1, 1998  
 APPLICATION NUMBER: 08/922,180  
 FILING DATE: September 2, 1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Friedman, Mark M.  
 REGISTRATION NUMBER: 33,883  
 REFERENCE/DOCKET NUMBER: 910/16  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 972-3-5625553  
 TELEFAX: 972-3-5625554  
 TELEX: <Unknown>  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 543  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 US-09-635-923-2

Query Match 37.2%; Score 1148.5; DB 4; Length 543;  
 Best Local Similarity 43.4%; Pred. No. 5.3e-110;  
 Matches 249; Conservative 82; Mismatches 190; Indels 53; Gaps 9;

18 PPACLPAGALYALILHLSSQAGDRRLPVDRRAAGLKEKTLILLDVSTKNPRTVNNEN 77  
 20 PLGLSPGAL-----PRPA-----QAQDVVDLFFTEGEPHLVSPS 55  
 78 FLISOLDPSIIHD-CWLDPLSKRLVTLARGSPAFIRFGGRKDPLOFQULRMPAKSRG 136  
 56 FLSTVIDNMLATDPFFLILGSPKLTARGLSPAYLRFPGTKIDFLIF---DPKEST 111  
 137 GGPDPYLYKNYEDDIVRSVALDKQGGCKIAQ-HEDVNLVIQREKXAMHLVLKKEQFSN 195  
 112 FEERSYWGQVNOPI-----CKYGSIPDVVEKRLRLEPVEQQLIREHYOK 158  
 196 TYSNLIILARSLDLYNPADESGGLHIFALNALRPNPNNSNSSALSILKYASKYNI 255  
 159 KFKSTYRSRSDVLYTPANCSDGLIFGLNALLETALQNMSSNAQCLLDYCSSKGYNI 218  
 256 SWELENEPNRYTRTHGRAVNSGOLGKDYIOLKSLLOPIRISRASLYGNIGRRKNVIA 315  
 219 SWELENEPNRSLFKKADIFINSQGEDYIQLHLKLRK-STFKNALTYGPDVGQRRKRYAK 277  
 316 LIDGMKAYAGSTVDAVYTHQCHIDRVRVYKMDFLKRLDLDLSQDIRIKQVNTYTPGK 375  
 278 MLKSTFLKAGGEYIDVYTHHYYLNGRTAIREDFLNPVDLDFISSQGVFQVVESTTRPGK 337  
 376 KIWLEGVVTSAGTNNISDSYAAQFLWMLNTLGLMANOGIDVIVIRHSFFDHGYNHLVDON 435

338 KWLIGETSSAYGGAPLLSDTFAGFWMLDKLGSAENGIDVNRQVFFGAGNYHLVDEN 397  
 436 FNPILPDYMLSLYKXLIQPKYLAHVAGLQRRKRGVIRDKLRIYACTNHHNNHYRG 495  
 398 FDPDPYMLSLFKLVGTGKYLMAVSQSKR-----KLVYLLHCTNTDPRYMEG 448  
 496 STTLFIINLHRSRKKIKIAGTRDKLVHGYLLQPYGEGLSKSVQNLGQPLVWYDQTL 555  
 449 DLTLYAINLHNTKYLRLPYPSNKKQVDRKYLRLPGFPGLLSKSVQNLGTLKWDQTL 508  
 556 PELKRPRLAGRTLVIPVWGMFYVKNVNLAC 589  
 509 PLMEKPLRPSGLGPAFYSFFVIRAKYAAAC 542

RESULT 8  
 US-09-487-716A-2  
 Sequence 2, Application US/09487716A  
 Patent No. 6475763  
 GENERAL INFORMATION:  
 APPLICANT: Macy Ayal-Herskovitz et al.  
 TITLE OF INVENTION: GENETICALLY MODIFIED CELLS AND METHODS FOR  
 EXPRESSING RECOMBINANT HEPARINASE  
 AND METHODS OF PURIFYING SAME  
 NUMBER OF SEQUENCES: 25  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Mark M. Friedman c/o Anthony Castorina  
 STREET: 2001 Jefferson Davis Highway, Suite 207  
 CITY: Arlington  
 STATE: Virginia  
 COUNTRY: United States of America  
 ZIP: 22202  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
 COMPUTER: PC  
 OPERATING SYSTEM: MS DOS version 6.2,  
 Windows version 3.11  
 SOFTWARE: Word for Windows version 2.0 converted to an ASCII file  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/487,716A  
 FILING DATE: 19-Jan-2000  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 09/071,618  
 FILING DATE: May 1, 1998  
 APPLICATION NUMBER: 09/071,739  
 FILING DATE: May 1, 1998  
 APPLICATION NUMBER: 08/922,180  
 FILING DATE: September 2, 1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Friedman, Mark M.  
 REGISTRATION NUMBER: 33,883  
 REFERENCE/DOCKET NUMBER: 910/16  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 972-3-5625553  
 TELEFAX: 972-3-5625554  
 TELEX: <Unknown>  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 543  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 US-09-487-716A-2

Query Match 37.2%; Score 1148.5; DB 4; Length 543;  
 Best Local Similarity 43.4%; Pred. No. 5.3e-110;  
 Matches 249; Conservative 82; Mismatches 190; Indels 53; Gaps 9;

18 PPACLPAGALYALILHLSSQAGDRRLPVDRRAAGLKEKTLILLDVSTKNPRTVNNEN 77

Fri May 14 13:01:55 2004

us-10-088-676-2.rat

Page 6

Db 20 PLGLSPGAL-----PRPA-----QAQDVVDLDFTOEPHLVSPS 55  
Qy 78 FLSLODPSIIHD-GMLDFLSKRLVTLARGLSPAFIRFGKRTDLOFONLNPAPSRG 136  
Db 56 FLSTVIDANLMTDPRFLILLSGPKLRTARGLSPAYIRFGSTKTIDFLIF----DPKKEST 111  
Qy 137 GPGPDYLYKNYEDDVRSDVALDKQKCKIAQ-HPDVMVLQREKAAQMHVLVLLKEQPSN 195  
Db 112 FEERSYQSQVNQDI-----CKYGSIPPDVEKRLRLEMPYQEOQL-LIREHYOK 158  
Qy 196 TYSNLILPARRSLDLYNFADCSGLHLIFALNALRRPNNSNASSALSLKYSASKXNYI 255  
Db 159 KFKNSTYSSSVDLVLYTFANCSSGLDLIFGLNALRTADLOMNSNAQLLDYSSSGYNI 218  
Qy 256 SWEIGNEPNRYRTMGRAVNGSQLGKDYIQLKSLQPIRYSRSLYGNIGRPRKNVIA 315  
Db 219 SWEIGNEPNRSFLKADIFINGSQLGEDYIQLHKLIRK-STFKNAKLYGPDVGQPRRRTAK 277  
Qy 316 LDGFMKVAAGSTDAVYMOHCYIDGRVYKVMDFLKTLLDLSQIRKIQVNTYTPGK 375  
Db 278 MLKSFLLKAGGEVIDSVTWHYTLNGRTATREDFLNPVDLIFISSVQKVFQVESTREPGK 337  
Qy 376 KIMLEGVTTTSGAGTNNLSDSYAAGFLMLNTLGMLANQIGIDVYIRHSFPDHGYNHLVDON 435  
Db 338 KVMLGSTSSAYGGAPLSDTFAAGFWMLDLGLSABGIEVYWRQVFFAGNHYLVDEN 397  
Qy 436 FNPDPYWLSTLYKRLIGPKVLAVHAGLQKRPGRVIRDKLRIYACHTNNHNNHYRG 495  
Db 398 FDLPLDYMSTLFLKLVGKVLMAVQSSKRR-----KLRVYLHCTVDNPRYREG 448  
Qy 496 SITFLINLHRSRKIKLAGTLRDLYVQYLLQPYGQGLSKSVQUNGOPLVWVDGTL 555  
Db 449 DLTVALNHNVTYKRLRPFPSNKQVDKYLIRPLGPHGLSKSVQUNGTLTKWVDGTL 508  
Qy 556 PELKPRPLRAGRTLVIPVTMGFYVVKVNAALAC 589  
Db 509 PLMEKPLRPGSSLGLPASFYSFVIRNAKVAAC 542

RESULT 9  
US-09-322-977-2  
Sequence 2, Application US/09322977  
Patent No. 6531129  
GENERAL INFORMATION:  
APPLICANT: Iris Pecker et al.  
TITLE OF INVENTION: HEPARANASE SPECIFIC MOLECULAR PROBES  
TITLE OF INVENTION: AND THEIR USE IN RESEARCH AND MEDICAL  
TITLE OF INVENTION: APPLICATIONS  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Mark M. Friedman c/o Anthony Castrorina  
STREET: 2001 Jefferson Davis Highway, Suite 207  
CITY: Arlington  
STATE: Virginia  
COUNTRY: United States of America  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
COMPUTER: Twinhead\* Slimote-890TX  
OPERATING SYSTEM: MS DOS version 6.2,  
OPERATING SYSTEM: Windows version 3.11  
SOFTWARE: Word for Windows version 2.0 converted to  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/322,977  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/922,180  
FILING DATE: September 2, 1997  
APPLICATION NUMBER: 09/071,739  
FILING DATE: May 1, 1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Friedman, Mark M.

REGISTRATION NUMBER: 33,883  
REFERENCE/DOCKET NUMBER: 910/21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 972-3-562553  
TELEFAX: 972-3-562554  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 543  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-322-977-2

Query Match 37.2% Score 1148.5; DB 4; Length 543;  
Best Local Similarity 43.4%; Pred. No. 5.3e-110;  
Matches 249; Conservative 82; Mismatches 190; Indels 53; Gaps 9;

Db 18 PPACIAPGALYALALHLHLSLSSQAQDRRPLFVDBAAGLKEKTLILLDVSTKNPVRTNEN 77  
Qy 20 PLGLSPGAL-----PRPA-----QAQDVVDLDFTOEPHLVSPS 55  
Db 78 FLSLODPSIIHD-GMLDFLSKRLVTLARGLSPAFIRFGKRTDLOFONLNPAPSRG 136  
Db 56 FLSTVIDANLMTDPRFLILLSGPKLRTARGLSPAYIRFGSTKTIDFLIF----DPKKEST 111  
Qy 137 GPGPDYLYKNYEDDVRSDVALDKQKCKIAQ-HPDVMVLQREKAAQMHVLVLLKEQPSN 195  
Db 112 FEERSYQSQVNQDI-----CKYGSIPPDVEKRLRLEMPYQEOQL-LIREHYOK 158  
Qy 196 TYSNLILPARRSLDLYNFADCSGLHLIFALNALRRPNNSNASSALSLKYSASKXNYI 255  
Db 159 KFKNSTYSSSVDLVLYTFANCSSGLDLIFGLNALRTADLOMNSNAQLLDYSSSGYNI 218  
Qy 256 SWEIGNEPNRYRTMGRAVNGSQLGKDYIQLKSLQPIRYSRSLYGNIGRPRKNVIA 315  
Db 219 SWEIGNEPNRSFLKADIFINGSQLGEDYIQLHKLIRK-STFKNAKLYGPDVGQPRRRTAK 277  
Qy 316 LDGFMKVAAGSTDAVYMOHCYIDGRVYKVMDFLKTLLDLSQIRKIQVNTYTPGK 375  
Db 278 MLKSFLLKAGGEVIDSVTWHYTLNGRTATREDFLNPVDLIFISSVQKVFQVESTREPGK 337  
Qy 376 KIMLEGVTTTSGAGTNNLSDSYAAGFLMLNTLGMLANQIGIDVYIRHSFPDHGYNHLVDON 435  
Db 338 KVMLGSTSSAYGGAPLSDTFAAGFWMLDLGLSABGIEVYWRQVFFAGNHYLVDEN 397  
Qy 436 FNPDPYWLSTLYKRLIGPKVLAVHAGLQKRPGRVIRDKLRIYACHTNNHNNHYRG 495  
Db 398 FDLPLDYMSTLFLKLVGKVLMAVQSSKRR-----KLRVYLHCTVDNPRYREG 448  
Qy 496 SITFLINLHRSRKIKLAGTLRDLYVQYLLQPYGQGLSKSVQUNGOPLVWVDGTL 555  
Db 449 DLTVALNHNVTYKRLRPFPSNKQVDKYLIRPLGPHGLSKSVQUNGTLTKWVDGTL 508  
Qy 556 PELKPRPLRAGRTLVIPVTMGFYVVKVNAALAC 589  
Db 509 PLMEKPLRPGSSLGLPASFYSFVIRNAKVAAC 542

RESULT 10  
US-09-186-200-1  
Sequence 1, Application US/09186200  
Patent No. 6562950  
GENERAL INFORMATION:  
APPLICANT: Perez, Tuvia et al.  
TITLE OF INVENTION: HEPARANASE ACTIVITY NEUTRALIZING ANTI-HEPARANASE MONOCLONAL ANTIBIC  
FILING DATE: 00/20441  
CURRENT APPLICATION NUMBER: US/09/186,200  
CURRENT FILING DATE: 2002-03-18  
NUMBER OF SEQ ID NOS: 1  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 1  
LENGTH: 543

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-186-200-1

```

Query Match	37.2%;	Score 1148.5;	DB 4;	Length 543;
Best Local Similarity	43.4%;	Pred. No. 5.3e-110;		
Matches 249;	Conservative 82;	Mismatches 190;	Indels 53;	Gaps 9.

[illegible]

RESULT 11  
US-09-435-739-10  
; Sequence 10, Application US/09435739

Query Match	37.2%;	Score 1148.5;	DB 4;	Length 543;
Best Local Similarity	43.4%;	Pred. No. 5.3e-110;		
Matches 249;	Conservative 82;	Mismatches 190;	Indels 53;	Gaps 9;

```

QY 18 P8ACIAPGALYIALLLHLSJSQAGDRRLPEVDBAAGKEKTLILDPVSTKNPRTYEN 77
Db 20 PUGJSPSGAL-----PRBA-----QADVDVDDLF7QCEPLHVSFS 55
QY 78 FLSIQDPSI1HD-GWLDFFLSKRLVTLARGISPAFLRFGGKRTDLOFONLENPAKSRG 136
Db 56 FLSVTIDALADPRFFLILSSPKRTIARLSPAYLRFPGTKTDLF-----DPKEST 111
QY 137 GPEPDYLLKXVEDDIYRSDVALDKOGCKIAQ-HPDYVLVIOQRKAQOMHVLILKEQFSN 195
Db 112 FEERSYQSQVNAQDI-----CKYSIPEDVEKRLRMPYQEOQ-LIREHYQK 156
QY 196 TYENLILTRASLDKLYNPAQCSGLHLFALNALBRPNMNSNSASVLSLKYSASKYNI 255
Db 159 KFNKSYSSSSVDVLYTFPANCSGLDILGMLNLRFTADLONSSMAQLLIDYCSKGYNI 218
QY 256 SWEIENGEPNNRYTMHGRAVNSQLOKXDYIOKSLQIRIYRSASLYGPNIGRPNKVIA 315
Db 219 SWEIENGEPNSFLKKADIFINSSQLOEEDYIOQHKLIRK-STFKNAKLYGPDVQGRRTAK 277
QY 316 LIDGPKVAVGSTVDVAVTMOCHYIDORVYKXWMDFLKTLILDLPLSLQIKQVNVVTPPK 375
Db 278 MLKSPFLKAGGEVIDSTWHYYLINRTRREDFLNPVDLIFISSVQVQFVGVESTBPGK 333
QY 376 KIMLEGVVTSAGTNNISDSYAAGFELMNTGMLANGIDVAVIRHSPFFDGYNHLVDON 435
Db 338 KVMLEGETSSAYGGAPLSDSTYPAAGFMMDLKGLSARBGIEVVRQVFFGAGNHLVDEN 399
QY 436 FNLPLDPYVLSLYKRLIGPKVLAVHVAQLOKRPBGRVIRDKRLIYACTNHHNNHYRG 499
Db 398 FDDLPLPYVLSLFLPKLVGTKVLMAVSQSKRR-----KLRYVLICTVTDNPRYEG 444
QY 496 SITLFLINLHRSKKYKLAAGTRDKLYHOYLLQPRQOGLSKSVQNLGQGLVAVDQGL 555
Db 449 DLTLYAINLHNVKYLRLPYPPSNKQVDCYLLRPLGPHGLSKSVQNLGTLTKWDDQGL 500
QY 556 PELKPRPLRAGRTLYVPVYMGVYVYKNNVNALAC 589
Db 509 PPLMEKPLRGGSSLGIPAFSYSPFYVIRAKVAAAC 542

```

```

1      RESULT 12
2      US-09-930-218-3
3      / Sequence 3, Application US/09930218
4      / Patent No. 6677137
5      / GENERAL INFORMATION:
6      / APPLICANT: goldsmidt, orit
7      / APPLICANT: pecker, itis
8      / APPLICANT: vlodavsky, israel
9      / APPLICANT: israel, michael
10     / TITLE OF INVENTION: AVIAN AND REPTILE DERIVED POLYNUCLEOTIDE ENCODING A POLYPEPTIDE H
11     / TITLE OF INVENTION: HEPARANASE ACTIVITY
12     / TITLE OF INVENTION: 01/22315
13     / FILE REFERENCE: 01/22315
14     / CURRENT APPLICATION NUMBER: US/09/930,218
15     / CURRENT FILING DATE: 2001-08-16
16     / PRIOR APPLICATION NUMBER: 09/666,390
17     / PRIOR FILING DATE: 2000-09-20
18     / NUMBER OF SEQ ID NOS: 16
19     / SOFTWARE: PatentIn version 3.1
20     / SEQ ID NO 3
21     / LENGTH: 543
22     / TYPE: PRT
23     / ORGANISM: Homo sapiens
24     / US-09-930-218-3

```

```

Query Match Similarity 37.2% Score 1148.5 - DB 4 Length 543;
Best Local Similarity 43.4% Pident No. 5.3e-110 Indels 53 Gaps 9;
Matches 249 Conservative 82 Mismatches 199
QY 18 PACACGATATATLHLSLSSQAGDRPPVYDRAAGAKERTLLIDYSTKNRPATVEN 77
DB 20 FPGSPSPAL-----FRP-----QAGGVVDVDFPFGFLHVPSP 55

```



```

QY 78 FLSLQDPSIIHD-GWLDPLSSKRLVTLARGLSPALFRFGKRTDFLOFONLRNPAKSRG 136
DB 56 FLSTVIDANLADPREFLILGSPKRLTLARGLSPALFRFGKRTDFLIF---DPEKST 111
QY 137 GPEPDYLYKNKEDDIYRSVDVALDKQKCKIAQ-HPDVMVLVOREKAQOHLVTLKEQFSN 195
DB 112 FEERSYQSQVNODI-----CKGSIIPDVEEKLRLMEWYEQOL-LLEHRYOK 158
QY 196 TYSNLLITARSJDKLVNFDSCGLHLIPALNARPNNSWNSSALSILKTSASKKXNI 255
DB 159 KFNKSTYSRSSVDLYTFANCSGLDLIFGLNALLRTRADLOMNSNQLLDYSSSGYNI 218
QY 256 SWEIGNEPNNSRTHGRANVGSQGDYIQLKSLQPIRIYSASLYGPNIGRPNKXIA 315
DB 219 SWEIGNEPNNSFLKKADIFINGSQLGEDIYQLHKLKLR-STFKNAKLYGPDVQPRKXIA 277
QY 316 LLDGFMKVASSTVDATWQHCYIDGRVYKVMDFLKTRELDLSDQIRKIQKVNTYTPGK 375
DB 278 MKKSPFKAGGEVIDSVTWHYINGRTRATEDFLNPDVLDIFISSVQKVFQVVESTRPGK 337
QY 376 KIVLEGVVTSAGCTNNLSYSYAGFLMNTLGMANOGIDVYIRSFEDHGYNHLVDON 435
DB 338 KIVLEGVTSAYYGGAPILSDTFAGFWMVDKLGISANMGLVVMRQVFFAGNYHLVDEN 397
QY 436 FNPIDYWLSTLYKRLIGPKYLAHVHAGLQKRPGRVIRDKLRIYAHCTNNHNYVRG 495
DB 398 FDPIDYWLSTLYKRLIGPKYLAHVHAGLQKRPGRVIRDKLRIYAHCTNNHNYVRG 448
QY 496 SITFLINLHRSRKKIKLAGTLADKLVHGYLLOPYGQEGLSKSVOLNCPVMVDGTL 555
DB 449 DLTLYALNHVNTKYLRPLPFPFSNKQVDKYLRLRGLPHGLSLSSVOLNGLTLKXVDQTL 508
QY 556 PELKRPRLRAGRTLVIPVTMGFYVYKVNALAC 589
DB 509 PLMEKPLRPGSSILGPAFSYSFVIRNNAKYAAC 542

RESULT 13
US-09-435-739-14
; Sequence 14, Application US/09435739
; Patent No. 6664105
; GENERAL INFORMATION:
; APPLICANT: Becker, Iris
; APPLICANT: Viodevsky, Israel
; APPLICANT: Feinstein, Elena
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY
; FILE REFERENCE: 00/20454
; CURRENT APPLICATION NUMBER: US/09/435,739
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 592
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-435-739-14

Query Match 37.2%; Score 1148.5; DB 4; Length 592;
Best Local Similarity 43.4%; Pred. No. 6.1e-110;
Matches 249; Conservative 82; Mismatches 190; Indels 53; Gaps 9;

```

```

QY 18 PPACIAPGALVYLLHLSLSQAGDRPLPVDRAAGLKEKTLILLDVSTKGPRTVYEN 77
DB 69 FLGPIPSFGL-----PRPA-----QAQVDVLDLFFTOEPFLHVS 104
QY 78 FLSLQDPSIIHD-GWLDPLSSKRLVTLARGLSPALFRFGKRTDFLOFONLRNPAKSRG 136
DB 56 FLSTVIDANLADPREFLILGSPKRLTLARGLSPALFRFGKRTDFLIF---DPEKST 111
QY 137 GPEPDYLYKNKEDDIYRSVDVALDKQKCKIAQ-HPDVMVLVOREKAQOHLVTLKEQFSN 195

```

```

DB 161 FEERSYQSQVNODI-----CKGSIIPDVEEKLRLMEWYEQOL-LLEHRYOK 207
QY 196 TYSNLLITARSJDKLVNFDSCGLHLIPALNARPNNSWNSSALSILKTSASKKXNI 255
DB 208 KFNKSTYSRSSVDLYTFANCSGLDLIFGLNALLRTRADLOMNSNQLLDYSSSGYNI 267
QY 256 SWEIGNEPNNSRTHGRANVGSQGDYIQLKSLQPIRIYSASLYGPNIGRPNKXIA 315
DB 268 SWEIGNEPNNSFLKKADIFINGSQLGEDIYQLHKLKLR-STFKNAKLYGPDVQPRKXIA 326
QY 316 LLDGFMKVASSTVDATWQHCYIDGRVYKVMDFLKTRELDLSDQIRKIQKVNTYTPGK 375
DB 327 MKKSPFKAGGEVIDSVTWHYINGRTRATEDFLNPDVLDIFISSVQKVFQVVESTRPGK 366
QY 376 KIVLEGVVTSAGCTNNLSYSYAGFLMNTLGMANOGIDVYIRSFEDHGYNHLVDON 435
DB 387 KIVLEGVTSAYYGGAPILSDTFAGFWMVDKLGISANMGLVVMRQVFFAGNYHLVDEN 446
QY 436 FNPIDYWLSTLYKRLIGPKYLAHVHAGLQKRPGRVIRDKLRIYAHCTNNHNYVRG 495
DB 447 FDPIDYWLSTLYKRLIGPKYLAHVHAGLQKRPGRVIRDKLRIYAHCTNNHNYVRG 497
QY 496 SITFLINLHRSRKKIKLAGTLADKLVHGYLLOPYGQEGLSKSVOLNCPVMVDGTL 555
DB 498 DLTLYALNHVNTKYLRPLPFPFSNKQVDKYLRLRGLPHGLSLSSVOLNGLTLKXVDQTL 557
QY 556 PELKRPRLRAGRTLVIPVTMGFYVYKVNALAC 589
DB 558 PLMEKPLRPGSSILGPAFSYSFVIRNNAKYAAC 591

RESULT 14
US-09-181-336-13
; Sequence 13, Application US/09181336A
; Patent No. 6242238
; GENERAL INFORMATION:
; APPLICANT: FREEMAN, Craig Geoffrey
; APPLICANT: BARISH, Richard
; APPLICANT: HAMDORF, Brenton James
; APPLICANT: HOLETT, Mark Darren
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE ENCODING MAMMALIAN
; FILE REFERENCE: 032505-004
; CURRENT APPLICATION NUMBER: US/09/181,336A
; CURRENT FILING DATE: 1998-10-28
; EARLIER APPLICATION NUMBER: AU PP0062
; EARLIER FILING DATE: 1997-10-28
; EARLIER APPLICATION NUMBER: AU PP0812
; EARLIER FILING DATE: 1997-12-09
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 543
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-181-336-13

Query Match 37.1%; Score 1144.5; DB 3; Length 543;
Best Local Similarity 43.2%; Pred. No. 1.4e-109;
Matches 248; Conservative 83; Mismatches 190; Indels 53; Gaps 9;

```

```

QY 18 PPACIAPGALVYLLHLSLSQAGDRPLPVDRAAGLKEKTLILLDVSTKGPRTVYEN 77
DB 20 FLGPIPSFGL-----PRPA-----QAQVDVLDLFFTOEPFLHVS 55
QY 78 FLSLQDPSIIHD-GWLDPLSSKRLVTLARGLSPALFRFGKRTDFLOFONLRNPAKSRG 136
DB 56 FLSTVIDANLADPREFLILGSPKRLTLARGLSPALFRFGKRTDFLIF---DPEKST 111
QY 137 GPEPDYLYKNKEDDIYRSVDVALDKQKCKIAQ-HPDVMVLVOREKAQOHLVTLKEQFSN 195
DB 112 FEERSYQSQVNODI-----CKGSIIPDVEEKLRLMEWYEQOL-LLEHRYOK 158

```

[illegible]

```

RESULT 15
US-09-601-777-2
; Sequence 2, Application US/09601777
; Patent No. 6461848
;
; GENERAL INFORMATION:
; APPLICANT: Nakajima, Motowo
; APPLICANT: Funakubo, Minako
; TITLE OF INVENTION: Human heparanase polypeptide and cDNA
; FILE REFERENCE: 30384A
; CURRENT APPLICATION NUMBER: US/09/601,777
; CURRENT FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;
; LENGTH: 568
;
; TYPE: PRT
;
; ORGANISM: Human
; US-09-601-777-2

```

Query Match	37.1%;	Score 1144.5;	DB 4;	Length 588;
Best Local Similarity	43.2%;	Pred. No. 1.6e-109;		
Matches 248;	Conservative 83;	Mismatches 150;	Indels 53;	Gaps 9;

```

Qy 1 PPAAGALVYLLUHLHLSOAGRRPLPDPBAAAGLKEKTLILDVSTGNPRATYEN 77
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 65 FLGLSLSGAL-----PRPA-----QAOVVDLDBFQEPHLHVSPS 100

Qy 78 FLSLQDPSLIHD-GWIDFLSKKLYTLARGSPAPLRFQGGKRDPLQFQNLRRPAAARG 136
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 101 FLVLTIDANATPERFLILGSEPKATTARGSPYLRGGKTDILF---DPKEST 156

Qy 137 GPGEDYLLKMYEDDIASVDALVDKQCKCIAQ-HPDVMVLQREKXAOMHLVLKEOFSN 195
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 157 PEEPSYQSQVNDI-----CKYGSIPDVEKRLBWPYEQOL-LIREHYQK 203

Qy 196 TYSNLIITAFSLDKLNFADCSGLHILPALNALRPNPNNSWSSALSILKYSASKTYN 255
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 204 KFKNSTYSRSASVYLYTFANCSGLDILFGNMLLRPADIQNNSSNAOALLDYCSSKGYN 263

Qy 256 SWEKGNPNPNRYRTMGHRAVNGSOLSDYIQLKSLQPIRKYRSAGIYSPNIGRPRVIA 315
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 264 SWEKGNPNPNFLKKADIFINGSQLGSDFIQJHKLARK-STFNALTYEPDVGQGRBRMTAK 322

Qy 316 LLDGFMKVAASSTDVATWQHCYIDGAVVAKVMDFLKRLLDLTISDQIRKIQVAVNTYIPGK 375

```

Db	323	MLMSFPAKAGEVIA DSTWHEHYLTNRKTAREDFLNPVLDLFISVQKVPQVVESTRPGK	382
Qy	376	KIMIEGCVTTSKAGCTNNLSDSYAAAGFLMLNTLQMLAANOGLDVRHSPFDGNNHLYDON	435
Db	383	KVVLGRTSSAYGGAGVLSDTFPAAGFMWLDKLGISARMGILVVRKQVFFGAGNTHLYDEN	442
Qy	436	FNELPDYMILSLYKRLIGEKVLAHVAGLQRKPRPGVRIRDKLRIYAHCTNNHNANTYRG	495
Db	443	FDELPYMYLSTLFKKLVGTKVLMAVSQSKR-----KLVYLIHCTNTDPRKEG	493
Qy	496	SITLFIINLHRSRKKIKIAGTLRDLVQVQYLLQPYQEGLSKSVQNLNGPVLVWVDDGTL	555
Db	494	DLTIYAINIHNATKYLRLEYPFSNNQVQKYLRLPGSPHGLSLKSVQNLNGTLKXVDDOTL	553
Qy	556	PELKPRLPAGRTLVLPPTWGFYVVKQVNNLAC	589
Db	554	PPLMEKPLRPGSSLGIPAPSYSFYIIRANXVAAAC	587

Search completed: May 6, 2004, 13:46:48  
Job time : 24 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2004, 13:45:47 ; Search time 49 Seconds  
(without alignments)  
3353.460 Million cell updates/sec

Title: US-10-088-676-2  
Perfect score: 3088  
Sequence: 1 MRVLCAPPEAMPSSNSRPPA.....FTVNGFYVKNVNLACRYR 592

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1140673 seqs, 27756755 residues

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3078	99.7	592	9	US-09-978-249-13
2	3071	99.4	592	14	US-10-177-245A-2
3	2740	88.7	534	9	US-09-880-262-2
4	2736	88.6	534	9	US-09-836-461-2
5	2736	88.6	534	14	US-10-177-245A-4
6	2446	79.2	480	9	US-09-978-249-7
7	2446	79.2	480	14	US-10-177-245A-6
8	2238	72.5	438	9	US-09-978-249-12
9	2238	72.5	438	9	US-09-930-218-2
10	1165	37.7	536	14	US-10-431-438-2
11	1163.5	37.7	535	9	US-09-930-218-1
12	1163.5	37.7	535	9	US-09-776-874A-44
13	1163.5	37.7	535	9	US-09-988-113-44
14	1163.5	37.7	535	14	US-10-341-582-44
15	1163.5	37.7	535	14	US-10-384-451-44

16	1163.5	37.7	535	14	US-10-431-438-1	Sequence 1, Appl
17	1163.5	37.7	535	14	US-10-384-450-44	Sequence 44, Appl
18	1163.5	37.7	535	15	US-10-371-218A-44	Sequence 44, Appl
19	1163.5	37.7	535	15	US-10-456-573-44	Sequence 44, Appl
20	1149	37.2	545	10	US-09-899-440-18	Sequence 18, Appl
21	1149	37.2	545	15	US-10-115-479-42	Sequence 42, Appl
22	1148.5	37.2	543	9	US-09-759-207-2	Sequence 2, Appl
23	1148.5	37.2	543	9	US-09-930-218-3	Sequence 3, Appl
24	1148.5	37.2	543	9	US-09-186-200-1	Sequence 1, Appl
25	1148.5	37.2	543	9	US-09-776-874A-10	Sequence 2, Appl
26	1148.5	37.2	543	9	US-09-944-602-2	Sequence 2, Appl
27	1148.5	37.2	543	9	US-09-322-977-2	Sequence 2, Appl
28	1148.5	37.2	543	9	US-09-988-113-10	Sequence 2, Appl
29	1148.5	37.2	543	12	US-10-676-079-2	Sequence 2, Appl
30	1148.5	37.2	543	14	US-10-137-351-2	Sequence 2, Appl
31	1148.5	37.2	543	14	US-10-341-582-10	Sequence 10, Appl
32	1148.5	37.2	543	14	US-10-384-451-10	Sequence 10, Appl
33	1148.5	37.2	543	14	US-10-431-438-3	Sequence 3, Appl
34	1148.5	37.2	543	14	US-10-368-044A-1	Sequence 1, Appl
35	1148.5	37.2	543	14	US-10-384-450-10	Sequence 10, Appl
36	1148.5	37.2	543	15	US-10-371-218A-10	Sequence 10, Appl
37	1148.5	37.2	543	15	US-10-456-573-10	Sequence 10, Appl
38	1148.5	37.2	543	9	US-09-776-874A-14	Sequence 14, Appl
39	1148.5	37.2	543	9	US-09-988-113-14	Sequence 14, Appl
40	1148.5	37.2	543	14	US-10-341-582-14	Sequence 14, Appl
41	1148.5	37.2	543	14	US-10-384-451-14	Sequence 14, Appl
42	1148.5	37.2	543	14	US-10-384-450-14	Sequence 14, Appl
43	1148.5	37.2	543	15	US-10-371-218A-14	Sequence 14, Appl
44	1148.5	37.2	543	15	US-10-456-573-14	Sequence 14, Appl
45	1144.5	37.1	554	14	US-10-314-683-2	Sequence 2, Appl

## ALIGNMENTS

RESULT 1  
US-09-978-249-13  
Sequence 13, Application US/0978249  
Patent No. US20020106780A1  
GENERAL INFORMATION:  
APPLICANT: Fissella, et al.  
TITLE OF INVENTION: Extracellular Matrix Polynucleotides, Polypeptides, and Antibodies  
FILE REFERENCE: PTO5491  
CURRENT FILING DATE: 2001-10-17  
CURRENT FILING DATE: 2001-10-17  
PRIOR FILING DATE: 2001-04-11  
PRIOR APPLICATION NUMBER: US/09/978, 249  
PRIOR FILING DATE: 2001-04-11  
PRIOR APPLICATION NUMBER: 60/198, 123  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 13  
LENGTH: 592  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-978-249-13

Query Match 99.7%; Score 3078; DB 9; Length 592;  
Best Local Similarity 99.7%; Pred. No. 4.3e-255;  
Matches 590; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Db 1  
QY 1 MRVLCAPPEAMPSSNSRPPALGALYLLHLSSSQGDRPPVPYRAAGLKEKTL 60  
1 MRVLCAPPEAMPSSNSRPPALGALYLLHLSSSQGDRPPVPYRAAGLKEKTL 60  
Db 61 ILLDVSTKNPVTVNENFLSLQDPSTIHDGMLFLSSKRLVTLARGSPALFEGKRT 120  
61 ILLDVSTKNPVTVNENFLSLQDPSTIHDGMLFLSSKRLVTLARGSPALFEGKRT 120  
QY 121 DFLQONTLRNPAKSGGPGPDYLLKNYEDD1VRSVDALDKQCKIAQHPVMTLQREK 180  
121 DFLQONTLRNPAKSGGPGPDYLLKNYEDD1VRSVDALDKQCKIAQHPVMTLQREK 180  
Db 122 DFLQONTLRNPAKSGGPGPDYLLKNYEDD1VRSVDALDKQCKIAQHPVMTLQREK 180  
122 DFLQONTLRNPAKSGGPGPDYLLKNYEDD1VRSVDALDKQCKIAQHPVMTLQREK 180

```

QY 181 AACMHLVLLKEQPSNTYSNLIITARSIDKYNFADCSGLHILPALNALRRPNNSWSSS 240
DB 181 AACMHLVLLKEQPSNTYSNLIITARSIDKYNFADCSGLHILPALNALRRPNNSWSSS 240
QY 241 AASLTKYSASKKYNIISWELGNEPNNRYTHGRAVNSQOLKDYIOUKSLQPIRISRAS 300
DB 241 AASLTKYSASKKYNIISWELGNEPNNRYTHGRAVNSQOLKDYIOUKSLQPIRISRAS 300
QY 301 LYGPNIQPRKXNTIALIDGFMKVAAGSTVDVAVWQHOCYIDGRVVKWDFLKTLDLTLSDQ 360
DB 301 LYGPNIQPRKXNTIALIDGFMKVAAGSTVDVAVWQHOCYIDGRVVKWDFLKTLDLTLSDQ 360
QY 361 IRIKQVNTYTPGKKIMLEGVTTSGAGTNNLSDSYAAGFLMLNTLGMLANQIDIVTR 420
DB 361 IRIKQVNTYTPGKKIMLEGVTTSGAGTNNLSDSYAAGFLMLNTLGMLANQIDIVTR 420
QY 421 HSFPDHGYNHLVDONENPLDPYMLSLYKRLIGPKYLAHVAGLQKPPRGVIRDKLRI 480
DB 421 HSFPDHGYNHLVDONENPLDPYMLSLYKRLIGPKYLAHVAGLQKPPRGVIRDKLRI 480
QY 481 YAHCTNHNHNYVSGITLFIINLHRSRKKIKLAGTIRDLVHQYLLQPYGQGLKSKSV 540
DB 481 YAHCTNHNHNYVSGITLFIINLHRSRKKIKLAGTIRDLVHQYLLQPYGQGLKSKSV 540
QY 541 QLNQGPLVMVDGTLPELKEPRLPAGRTLVIPTVWGFVYVKNVNALACRYR 592
DB 541 QLNQGPLVMVDGTLPELKEPRLPAGRTLVIPTVWGFVYVKNVNALACRYR 592

```

RESULT 2  
US-10-177-245A-2  
Sequence 2, Application US/10177245A  
Publication No. US20030083254A1

```

GENERAL INFORMATION:
APPLICANT: McKenzie, Edward Alexander
APPLICANT: Stamps, Aladeair Craig
APPLICANT: Terrett, Jonathan Alexander
APPLICANT: Tyson, Kerry Louise
TITLE OF INVENTION: Substances
FILE REFERENCE: 2543-1-027
CURRENT APPLICATION NUMBER: US/10/177,245A
CURRENT FILING DATE: 2002-08-29
PRIOR APPLICATION NUMBER: PCT/GB00/04963
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: UK 0008713.0
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: UK 9930392.7
PRIOR FILING DATE: 1999-12-22
NUMBER OF SEQ ID NOS: 23
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 592
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(592)
OTHER INFORMATION: Xaa = Any Amino Acid
US-10-177-245A-2

```

```

Query Match 99.4%; Score 3071; DB 14; Length 592;
Best Local Similarity 99.5%; Pred. No. 2, 1e-294;
Matches 589; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 MRVLCAPFAPMSSSRPPACIAPGALYIALLLHLSLSSQAGDRRPLFVDRRAAGLKEKTL 60
DB 1 MRVLCAPFAPMSSSRPPACIAPGALYIALLLHLSLSSQAGDRRPLFVDRRAAGLKEKTL 60
QY 61 ILLDVSTKNPRTVNENFLSLQLDPSIITHDGMDFLSKRLVTLARGISPAFLRFGKRT 120
DB 61 ILLDVSTKNPRTVNENFLSLQLDPSIITHDGMDFLSKRLVTLARGISPAFLRFGKRT 120
QY 121 DFLQFQNLNRPKAKSGGPGPDYLLKNYEDDIYRSVDALDKQKGCIAQHPDVMVLQREK 180
DB 121 DFLQFQNLNRPKAKSGGPGPDYLLKNYEDDIYRSVDALDKQKGCIAQHPDVMVLQREK 180

```

```

DB 121 DFLQFQNLNRPKAKSGGPGPDYLLKNYEDDIYRSVDALDKQKGCIAQHPDVMVLQREK 180
QY 181 AACMHLVLLKEQPSNTYSNLIITARSIDKYNFADCSGLHILPALNALRRPNNSWSSS 240
DB 181 AACMHLVLLKEQPSNTYSNLIITARSIDKYNFADCSGLHILPALNALRRPNNSWSSS 240
QY 241 AASLTKYSASKKYNIISWELGNEPNNRYTHGRAVNSQOLKDYIOUKSLQPIRISRAS 300
DB 241 AASLTKYSASKKYNIISWELGNEPNNRYTHGRAVNSQOLKDYIOUKSLQPIRISRAS 300
QY 301 LYGNIRGPRKXNTIALIDGFMKVAAGSTVDVAVWQHOCYIDGRVVKWDFLKTLDLTLSDQ 360
DB 301 LYGNIRGPRKXNTIALIDGFMKVAAGSTVDVAVWQHOCYIDGRVVKWDFLKTLDLTLSDQ 360
QY 361 IRIKQVNTYTPGKKIMLEGVTTSGAGTNNLSDSYAAGFLMLNTLGMLANQIDIVTR 420
DB 361 IRIKQVNTYTPGKKIMLEGVTTSGAGTNNLSDSYAAGFLMLNTLGMLANQIDIVTR 420
QY 421 HSFPDHGYNHLVDONENPLDPYMLSLYKRLIGPKYLAHVAGLQKPPRGVIRDKLRI 480
DB 421 HSFPDHGYNHLVDONENPLDPYMLSLYKRLIGPKYLAHVAGLQKPPRGVIRDKLRI 480
QY 481 YAHCTNHNHNYVSGITLFIINLHRSRKKIKLAGTIRDLVHQYLLQPYGQGLKSKSV 540
DB 481 YAHCTNHNHNYVSGITLFIINLHRSRKKIKLAGTIRDLVHQYLLQPYGQGLKSKSV 540
QY 541 QLNQGPLVMVDGTLPELKEPRLPAGRTLVIPTVWGFVYVKNVNALACRYR 592
DB 541 QLNQGPLVMVDGTLPELKEPRLPAGRTLVIPTVWGFVYVKNVNALACRYR 592

```

RESULT 3  
US-09-880-262-2  
Sequence 2, Application US/09880262  
Patent No. US20020137907A1

```

GENERAL INFORMATION:
APPLICANT: LAMWINE, CHRISTOPHER DONALD
APPLICANT: SOUTMAN, CHRISTOPHER GEOFFREY CARSON
APPLICANT: RANGE, KIM
APPLICANT: HAYES, PHILIP DAVID
TITLE OF INVENTION: NOVEL COMPOUNDS
FILE REFERENCE: GP-30231
CURRENT APPLICATION NUMBER: US/09/880,262
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: UK 0014447.7
PRIOR FILING DATE: 2000-06-13
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 534
TYPE: PRT
ORGANISM: HOMO SAPIENS
US-09-880-262-2

```

```

Query Match 88.7%; Score 2740; DB 9; Length 534;
Best Local Similarity 90.0%; Pred. No. 1e-261;
Matches 533; Conservative 0; Mismatches 1; Indels 58; Gaps 1;
QY 1 MRVLCAPFAPMSSSRPPACIAPGALYIALLLHLSLSSQAGDRRPLFVDRRAAGLKEKTL 60
DB 1 MRVLCAPFAPMSSSRPPACIAPGALYIALLLHLSLSSQAGDRRPLFVDRRAAGLKEKTL 60
QY 61 ILLDVSTKNPRTVNENFLSLQLDPSIITHDGMDFLSKRLVTLARGISPAFLRFGKRT 120
DB 61 ILLDVSTKNPRTVNENFLSLQLDPSIITHDGMDFLSKRLVTLARGISPAFLRFGKRT 120
QY 121 DFLQFQNLNRPKAKSGGPGPDYLLKNYEDDIYRSVDALDKQKGCIAQHPDVMVLQREK 180
DB 121 DFLQFQNLNRPKAKSGGPGPDYLLKNYEDDIYRSVDALDKQKGCIAQHPDVMVLQREK 180
QY 181 AACMHLVLLKEQPSNTYSNLIITARSIDKYNFADCSGLHILPALNALRRPNNSWSSS 240
DB 181 AACMHLVLLKEQPSNTYSNLIITARSIDKYNFADCSGLHILPALNALRRPNNSWSSS 240

```

```

Db      181 AAOMHLVLLKEQFSNTYSNLIIT----- 203
Qy      241 ALSILKYSASKKYNISWELGNEPNRYRTWEGRAVNSQOLGKDYIOLKSLLOPIRISRAS 300
Db      204 -----EPNNRYRTWEGRAVNSQOLGKDYIOLKSLLOPIRISRAS 242
Qy      301 LYGNIGRPRKNVIALLDGFMKVAGSTVDVATWQHCHTIDRVRVKNWDFLKTLLDLSQ 360
Db      243 LYGNIGRPRKNVIALLDGFMKVAGSTVDVATWQHCHTIDRVRVKNWDFLKTLLDLSQ 302
Qy      361 IRIKQKVVNTYTPGKKTMLBEGVTTTSAGTNNLSDSAAGFLWMLTGLMANGIDIVIR 420
Db      303 IRIKQKVVNTYTPGKKTMLBEGVTTTSAGTNNLSDSAAGFLWMLTGLMANGIDIVIR 362
Qy      421 HSFPDHGYNHLVDQNFNLPDYWLSLYKRLIGKVLAVHAGLQKRPGRVIRDKLRI 480
Db      363 HSFPDHGYNHLVDQNFNLPDYWLSLYKRLIGKVLAVHAGLQKRPGRVIRDKLRI 422
Qy      481 YAHCTNHNNHYVRSITLFTIHLRSRKKIKLAGTLDKLVHGYLLQPYQOGLKSKSV 540
Db      423 YAHCTNHNNHYVRSITLFTIHLRSRKKIKLAGTLDKLVHGYLLQPYQOGLKSKSV 482
Qy      541 QLNQGPLVWVDGTLPELKRPRPLRAGRTLVIPVTMGFFVVKNNALACRYR 592
Db      483 QLNQGPLVWVDGTLPELKRPRPLRAGRTLVIPVTMGFFVVKNNALACRYR 534

RESULT 4
US-09-836-461-2
; Sequence 2, Application US/09836461
; Patent No. US20020064853A1
; GENERAL INFORMATION:
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Hejnikscs, Robert L.
; TITLE OF INVENTION: Heparanase II, A No. US20020064853A1e1 Human Heparanase Paralog
; FILE REFERENCE: heparanase II
; CURRENT APPLICATION NUMBER: US/09/836,461
; CURRENT FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-836-461-2

Query Match      88.6%; Score 2736; DB 9; Length 534;
Best Local Similarity 89.9%; Pred. No. 2,6e-261;
Matches 532; Conservative 1; Mismatches 1; Indels 58; Gaps 1;

Qy      1 MRVLCAPPEAMPSSNSRPPACIAPGALYIALLLHLSLSSQAGDRRPLPVDRAGLKEKTL 60
Db      1 MRVLCAPPEAMPSSNSRPPACIAPGALYIALLLHLSLSSQAGDRRPLPVDRAGLKEKTL 60
Qy      61 ILLDVSTKNPRTVNVENFLSLQDPSIITHDGLDPLSKRLVTLARGISPAFLRGKRT 120
Db      61 ILLDVSTKNPRTVNVENFLSLQDPSIITHDGLDPLSKRLVTLARGISPAFLRGKRT 120
Qy      121 DFLQFONLRNPAKSRGPGPDYLLKNYEDDIYRSVDVADLKQKGCIAQHPDVMVLQREK 180
Db      121 DFLQFONLRNPAKSRGPGPDYLLKNYEDDIYRSVDVADLKQKGCIAQHPDVMVLQREK 180
Qy      181 AAQMHVLVLLKEQFSNTYSNLIITARSIDLKYNFADCSGLHLIFALNALRRPNNSWSSS 240
Db      181 AAQMHVLVLLKEQFSNTYSNLIITARSIDLKYNFADCSGLHLIFALNALRRPNNSWSSS 240
Qy      241 ALSILKYSASKKYNISWELGNEPNRYRTWEGRAVNSQOLGKDYIOLKSLLOPIRISRAS 300
Db      204 -----EPNNRYRTWEGRAVNSQOLGKDYIOLKSLLOPIRISRAS 242
Qy      301 LYGNIGRPRKNVIALLDGFMKVAGSTVDVATWQHCHTIDRVRVKNWDFLKTLLDLSQ 360
Db      243 LYGNIGRPRKNVIALLDGFMKVAGSTVDVATWQHCHTIDRVRVKNWDFLKTLLDLSQ 302

```

```

Qy      361 IRIKQKVVNTYTPGKKTMLBEGVTTTSAGTNNLSDSAAGFLWMLTGLMANGIDIVIR 420
Db      303 IRIKQKVVNTYTPGKKTMLBEGVTTTSAGTNNLSDSAAGFLWMLTGLMANGIDIVIR 362
Qy      421 HSFPDHGYNHLVDQNFNLPDYWLSLYKRLIGKVLAVHAGLQKRPGRVIRDKLRI 480
Db      363 HSFPDHGYNHLVDQNFNLPDYWLSLYKRLIGKVLAVHAGLQKRPGRVIRDKLRI 422
Qy      481 YAHCTNHNNHYVRSITLFTIHLRSRKKIKLAGTLDKLVHGYLLQPYQOGLKSKSV 540
Db      423 YAHCTNHNNHYVRSITLFTIHLRSRKKIKLAGTLDKLVHGYLLQPYQOGLKSKSV 482
Qy      541 QLNQGPLVWVDGTLPELKRPRPLRAGRTLVIPVTMGFFVVKNNALACRYR 592
Db      483 QLNQGPLVWVDGTLPELKRPRPLRAGRTLVIPVTMGFFVVKNNALACRYR 534

```

```

RESULT 5
US-10-177-245A-4
; Sequence 4, Application US/10177245A
; Publication No. US20030083254A1
; GENERAL INFORMATION:
; APPLICANT: McKenzie, Edward Alexander
; APPLICANT: Stamps, Alasdair Craig
; APPLICANT: Terrett, Jonathan Alexander
; APPLICANT: Tyson, Kerry Louise
; TITLE OF INVENTION: Substances
; FILE REFERENCE: 2543-1-027
; CURRENT APPLICATION NUMBER: US/10/177,245A
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: PCT/GB00/04963
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: UK 0008713.0
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: UK 9930392.7
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-177-245A-4

```

```

Query Match      88.6%; Score 2736; DB 14; Length 534;
Best Local Similarity 89.9%; Pred. No. 2,6e-261;
Matches 532; Conservative 1; Mismatches 1; Indels 58; Gaps 1;

Qy      1 MRVLCAPPEAMPSSNSRPPACIAPGALYIALLLHLSLSSQAGDRRPLPVDRAGLKEKTL 60
Db      1 MRVLCAPPEAMPSSNSRPPACIAPGALYIALLLHLSLSSQAGDRRPLPVDRAGLKEKTL 60
Qy      61 ILLDVSTKNPRTVNVENFLSLQDPSIITHDGLDPLSKRLVTLARGISPAFLRGKRT 120
Db      61 ILLDVSTKNPRTVNVENFLSLQDPSIITHDGLDPLSKRLVTLARGISPAFLRGKRT 120
Qy      121 DFLQFONLRNPAKSRGPGPDYLLKNYEDDIYRSVDVADLKQKGCIAQHPDVMVLQREK 180
Db      121 DFLQFONLRNPAKSRGPGPDYLLKNYEDDIYRSVDVADLKQKGCIAQHPDVMVLQREK 180
Qy      181 AAQMHVLVLLKEQFSNTYSNLIITARSIDLKYNFADCSGLHLIFALNALRRPNNSWSSS 240
Db      181 AAQMHVLVLLKEQFSNTYSNLIITARSIDLKYNFADCSGLHLIFALNALRRPNNSWSSS 240
Qy      241 ALSILKYSASKKYNISWELGNEPNRYRTWEGRAVNSQOLGKDYIOLKSLLOPIRISRAS 300
Db      204 -----EPNNRYRTWEGRAVNSQOLGKDYIOLKSLLOPIRISRAS 242
Qy      301 LYGNIGRPRKNVIALLDGFMKVAGSTVDVATWQHCHTIDRVRVKNWDFLKTLLDLSQ 360
Db      243 LYGNIGRPRKNVIALLDGFMKVAGSTVDVATWQHCHTIDRVRVKNWDFLKTLLDLSQ 302

```

```

QY 361 IRKIQKVNTYTTGGKIMLEGVTTTSAGTNNISDYAAGFLMNTLGMANGIDIVIR 420
Db 303 IRKIQKVNTYTTGGKIMLEGVTTTSAGTNNISDYAAGFLMNTLGMANGIDIVIR 362
QY 421 HSFEDHGYNHLVDQNFNPLDPYMLSILYKELIGPKYLAHVAGLQKRPGRVIRDKLRI 480
Db 363 HSFEDHGYNHLVDQNFNPLDPYMLSILYKELIGPKYLAHVAGLQKRPGRVIRDKLRI 422
QY 481 YAHCTNNHNNHYVGSITLFTINLHRSRKIKLAGTIRDLVHOYLQPYGGBGLSKSV 540
Db 423 YAHCTNNHNNHYVGSITLFTINLHRSRKIKLAGTIRDLVHOYLQPYGGBGLSKSV 482
QY 541 QUNGQPLVMVDDGTLPELKRPRLRAGRTLVIPVTMGFFVYKVNALACRYR 592
Db 483 QUNGQPLVMVDDGTLPELKRPRLRAGRTLVIPVTMGFFVYKVNALACRYR 534

```

# RESULT 6

```

US-09-978-249-7
; Sequence 7, Application US/09978249
; Patent No. US20020106780A1
; GENERAL INFORMATION:
; APPLICANT: Fiscella, et al.
; TITLE OF INVENTION: Extracellular Matrix Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: P05491
; CURRENT APPLICATION NUMBER: US/09/978, 249
; PRIOR FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: PCT/US01/11643
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/198,123
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-978-249-7

```

Query Match 79.2%; Score 2446; DB 9; Length 480;  
 Best Local Similarity 80.9%; Pred. No. 1.1e-232;  
 Matches 479; Conservative 1; Mismatches 0; Indels 112; Gaps 1;

```

QY 1 MRVLCAPFEMPSSNSRPPACIAPGALYIALHLHLSLSSQAGDRRPLVDRRAAGLKEKTL 60
Db 1 MRVLCAPFEMPSSNSRPPACIAPGALYIALHLHLSLSSQAGDRRPLVDRRAAGLKEKTL 60
QY 61 ILLDVSTKNPRTVVENFLSLQDPSIHDGMLDFLSKRLVTLARGISPAFLRFGKRT 120
Db 61 ILLDVSTKNPRTVVENFLSLQDPSIHDGMLDFLSKRLVTLARGISPAFLRFGKRT 120
QY 121 DFLQFQNLNPAKSRGPGPDYLLKNYEDIVRSVDVALDKQCKIAQHPDVMVLQREK 180
Db 121 DFLQFQNLNPAKSRGPGPDYLLKNYEDIVRSVDVALDKQCKIAQHPDVMVLQREK 180
QY 149 121 DFLQFQNLNPAKSRGPGPDYLLKNYEDIVRSVDVALDKQCKIAQHPDVMVLQREK 149
Db 149 121 DFLQFQNLNPAKSRGPGPDYLLKNYEDIVRSVDVALDKQCKIAQHPDVMVLQREK 149
QY 181 AAQWMLVLLKQPSNTYSNLLITARSCLKYNFADCSGLHLIFALNALRRPNNSWSSS 240
Db 181 AAQWMLVLLKQPSNTYSNLLITARSCLKYNFADCSGLHLIFALNALRRPNNSWSSS 240
QY 149 150 -----EPNNYRTMHGRAVNGSLQGDYIQLKSLQPIRIYSRAS 149
Db 149 150 -----EPNNYRTMHGRAVNGSLQGDYIQLKSLQPIRIYSRAS 149
QY 241 ALSLKYSASKKYNISWELGNEPNNYRTMHGRAVNGSLQGDYIQLKSLQPIRIYSRAS 300
Db 241 ALSLKYSASKKYNISWELGNEPNNYRTMHGRAVNGSLQGDYIQLKSLQPIRIYSRAS 300
QY 150 -----EPNNYRTMHGRAVNGSLQGDYIQLKSLQPIRIYSRAS 188
Db 150 -----EPNNYRTMHGRAVNGSLQGDYIQLKSLQPIRIYSRAS 188
QY 301 LYGPNIQPRKRVIALLDGFMKVASGYDAVTWQHCYIDGRVYKMDFLKTRLLDITLSDQ 360
Db 301 LYGPNIQPRKRVIALLDGFMKVASGYDAVTWQHCYIDGRVYKMDFLKTRLLDITLSDQ 360
QY 189 LYGPNIQPRKRVIALLDGFMKVASGYDAVTWQHCYIDGRVYKMDFLKTRLLDITLSDQ 248
Db 189 LYGPNIQPRKRVIALLDGFMKVASGYDAVTWQHCYIDGRVYKMDFLKTRLLDITLSDQ 248
QY 361 IRKIQKVNTYTTGGKIMLEGVTTTSAGTNNISDYAAGFLMNTLGMANGIDIVIR 420
Db 361 IRKIQKVNTYTTGGKIMLEGVTTTSAGTNNISDYAAGFLMNTLGMANGIDIVIR 420
QY 249 IRKIQKVNTYTTGGKIMLEGVTTTSAGTNNISDYAAGFLMNTLGMANGIDIVIR 308
Db 249 IRKIQKVNTYTTGGKIMLEGVTTTSAGTNNISDYAAGFLMNTLGMANGIDIVIR 308
QY 421 HSFEDHGYNHLVDQNFNPLDPYMLSILYKELIGPKYLAHVAGLQKRPGRVIRDKLRI 480
Db 421 HSFEDHGYNHLVDQNFNPLDPYMLSILYKELIGPKYLAHVAGLQKRPGRVIRDKLRI 480

```

```

Db 309 HSFEDHGYNHLVDQNFNPLDPYMLSILYKELIGPKYLAHVAGLQKRPGRVIRDKLRI 368
QY 481 YAHCTNNHNNHYVGSITLFTINLHRSRKIKLAGTIRDLVHOYLQPYGGBGLSKSV 540
Db 369 YAHCTNNHNNHYVGSITLFTINLHRSRKIKLAGTIRDLVHOYLQPYGGBGLSKSV 428
QY 541 QUNGQPLVMVDDGTLPELKRPRLRAGRTLVIPVTMGFFVYKVNALACRYR 592
Db 429 QUNGQPLVMVDDGTLPELKRPRLRAGRTLVIPVTMGFFVYKVNALACRYR 480

```

# RESULT 7

```

US-10-177-245A-6
; Sequence 6, Application US/10177245A
; Publication No. US20030083254A1
; GENERAL INFORMATION:
; APPLICANT: McKenzie, Edward Alexander
; APPLICANT: Stamps, Alasdair Craig
; APPLICANT: Terrett, Jonathan Alexander
; APPLICANT: Tyson, Kerry Louise
; TITLE OF INVENTION: Substances
; FILE REFERENCE: 2543-1-027
; CURRENT APPLICATION NUMBER: US/10/177, 245A
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: PCT/GB00/04963
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: UK 0008713.0
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: UK 9930392.7
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-245A-6

```

Query Match 79.2%; Score 2446; DB 14; Length 480;  
 Best Local Similarity 80.9%; Pred. No. 1.1e-232;  
 Matches 479; Conservative 1; Mismatches 0; Indels 112; Gaps 1;

```

QY 1 MRVLCAPFEMPSSNSRPPACIAPGALYIALHLHLSLSSQAGDRRPLVDRRAAGLKEKTL 60
Db 1 MRVLCAPFEMPSSNSRPPACIAPGALYIALHLHLSLSSQAGDRRPLVDRRAAGLKEKTL 60
QY 61 ILLDVSTKNPRTVVENFLSLQDPSIHDGMLDFLSKRLVTLARGISPAFLRFGKRT 120
Db 61 ILLDVSTKNPRTVVENFLSLQDPSIHDGMLDFLSKRLVTLARGISPAFLRFGKRT 120
QY 121 DFLQFQNLNPAKSRGPGPDYLLKNYEDIVRSVDVALDKQCKIAQHPDVMVLQREK 180
Db 121 DFLQFQNLNPAKSRGPGPDYLLKNYEDIVRSVDVALDKQCKIAQHPDVMVLQREK 180
QY 149 121 DFLQFQNLNPAKSRGPGPDYLLKNYEDIVRSVDVALDKQCKIAQHPDVMVLQREK 149
Db 149 121 DFLQFQNLNPAKSRGPGPDYLLKNYEDIVRSVDVALDKQCKIAQHPDVMVLQREK 149
QY 181 AAQWMLVLLKQPSNTYSNLLITARSCLKYNFADCSGLHLIFALNALRRPNNSWSSS 240
Db 181 AAQWMLVLLKQPSNTYSNLLITARSCLKYNFADCSGLHLIFALNALRRPNNSWSSS 240
QY 149 150 -----EPNNYRTMHGRAVNGSLQGDYIQLKSLQPIRIYSRAS 149
Db 149 150 -----EPNNYRTMHGRAVNGSLQGDYIQLKSLQPIRIYSRAS 149
QY 241 ALSLKYSASKKYNISWELGNEPNNYRTMHGRAVNGSLQGDYIQLKSLQPIRIYSRAS 300
Db 241 ALSLKYSASKKYNISWELGNEPNNYRTMHGRAVNGSLQGDYIQLKSLQPIRIYSRAS 300
QY 150 -----EPNNYRTMHGRAVNGSLQGDYIQLKSLQPIRIYSRAS 188
Db 150 -----EPNNYRTMHGRAVNGSLQGDYIQLKSLQPIRIYSRAS 188
QY 301 LYGPNIQPRKRVIALLDGFMKVASGYDAVTWQHCYIDGRVYKMDFLKTRLLDITLSDQ 360
Db 301 LYGPNIQPRKRVIALLDGFMKVASGYDAVTWQHCYIDGRVYKMDFLKTRLLDITLSDQ 360
QY 189 LYGPNIQPRKRVIALLDGFMKVASGYDAVTWQHCYIDGRVYKMDFLKTRLLDITLSDQ 248
Db 189 LYGPNIQPRKRVIALLDGFMKVASGYDAVTWQHCYIDGRVYKMDFLKTRLLDITLSDQ 248
QY 361 IRKIQKVNTYTTGGKIMLEGVTTTSAGTNNISDYAAGFLMNTLGMANGIDIVIR 420
Db 361 IRKIQKVNTYTTGGKIMLEGVTTTSAGTNNISDYAAGFLMNTLGMANGIDIVIR 420
QY 249 IRKIQKVNTYTTGGKIMLEGVTTTSAGTNNISDYAAGFLMNTLGMANGIDIVIR 308
Db 249 IRKIQKVNTYTTGGKIMLEGVTTTSAGTNNISDYAAGFLMNTLGMANGIDIVIR 308
QY 421 HSFEDHGYNHLVDQNFNPLDPYMLSILYKELIGPKYLAHVAGLQKRPGRVIRDKLRI 480
Db 309 HSFEDHGYNHLVDQNFNPLDPYMLSILYKELIGPKYLAHVAGLQKRPGRVIRDKLRI 368

```



QY 481 YACHTNNHNNYVRSITLFIINLHRSKKIKLAGTLDKLVHGYLLQPYGEGLSKSV 540  
 DB 369 YACHTNNHNNYVRSITLFIINLHRSKKIKLAGTLDKLVHGYLLQPYGEGLSKSV 428  
 QY 541 QUNGOPLVMDGTLPELKRPLRAGTLYIPVTMGFVYVKNVNLACRYR 592  
 DB 429 QUNGOPLVMDGTLPELKRPLRAGTLYIPVTMGFVYVKNVNLACRYR 480

## RESULT 8

US-09-978-249-12  
 ; Sequence 12, Application US/0978249  
 ; Patent No. US20020106780A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Fiscella, et al.  
 ; TITLE OF INVENTION: Extracellular Matrix Polynucleotides, Polypeptides, and Antibodies  
 ; FILE REFERENCE: P054P1  
 ; CURRENT APPLICATION NUMBER: US/09/978,249  
 ; CURRENT FILING DATE: 2001-10-17  
 ; PRIOR APPLICATION NUMBER: PCT/US01/11643  
 ; PRIOR FILING DATE: 2001-04-11  
 ; PRIOR APPLICATION NUMBER: 60/198,123  
 ; PRIOR FILING DATE: 2000-04-18  
 ; NUMBER OF SEQ ID NOS: 16  
 ; SOFTWARE: Patent In Ver. 2.0  
 ; SEQ ID NO 12  
 ; LENGTH: 439  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-978-249-12

Query Match 72.5%; Score 2238; DB 9; Length 439;  
 Best Local Similarity 79.5%; Pred. No. 3,66-212; Indels 112; Gaps 1;  
 Matches 439; Conservative 1; Mismatches 0;

QY 42 GDRRLPVDRAGLKEKTLILLDVSTKNPVRYVNEFLSLQDPSIHDGMLDPLSSKRL 101  
 DB 1 GDRRLPVDRAGLKEKTLILLDVSTKNPVRYVNEFLSLQDPSIHDGMLDPLSSKRL 60  
 QY 102 VTLARGLSPAFLRFGKRTDPLQFONLRNPAKSGGPGDYLYKNEEDIVRSDVALDKQ 161  
 DB 61 VTLARGLSPAFLRFGKRTDPLQFONLRNPAKSGGPGDYLYKNEEDIVRSDVALDKQ 108  
 QY 162 KCKGKIAHPVMTVLOREKAAQWHLVILKEQFNTYSNLTITARSIDKLVPADCSGLHL 221  
 DB 109 ----- 108  
 QY 222 IFALNALRPNPNSWSSSALSILKYSASKYNIWELGNEPNRYTMHGRAVNGSOLG 281  
 DB 109 ----- EPNRYTMHGRAVNGSOLG 128  
 QY 282 DYIOLKSLQPIRIRYSASLYGNIGRPRKNVIALLDGFMKVASIVDAVWQHCYIDGR 341  
 DB 129 DYIOLKSLQPIRIRYSASLYGNIGRPRKNVIALLDGFMKVASIVDAVWQHCYIDGR 188  
 QY 342 VVKYMDLKTLLDLSDOIIRKIOKVNTYTPGKKILLEGVTTISAGTNNLSYSYAGF 401  
 DB 189 VVKYMDLKTLLDLSDOIIRKIOKVNTYTPGKKILLEGVTTISAGTNNLSYSYAGF 248  
 QY 402 LMLNTLGMLANOGIDVYIRHSFDFHGYNHLVDQNFPLPDYWLSTLYKRLIGPKYLAHV 461  
 DB 249 LMLNTLGMLANOGIDVYIRHSFDFHGYNHLVDQNFPLPDYWLSTLYKRLIGPKYLAHV 308  
 QY 462 AGLOKRPGRVITDKLRIYACHTNNHNNYVRSITLFIINLHRSKKIKLAGTLDKLV 521  
 DB 309 AGLOKRPGRVITDKLRIYACHTNNHNNYVRSITLFIINLHRSKKIKLAGTLDKLV 368  
 QY 522 VHOYILOPYGEGLSKSVQUNGOPLVMDGTLPELKRPLRAGTLYIPVTMGFVYV 581  
 DB 369 VHOYILOPYGEGLSKSVQUNGOPLVMDGTLPELKRPLRAGTLYIPVTMGFVYV 428  
 QY 582 KVNNALACRYR 592

DB 429 KVNNALACRYR 439

## RESULT 9

US-09-930-218-2  
 ; Sequence 2, Application US/09930218  
 ; Patent No. US20020034810A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: goldsmith, oric  
 ; APPLICANT: pecker, iris  
 ; APPLICANT: vlodavsky, israel  
 ; TITLE OF INVENTION: AVIAN AND REPTILE DERIVED POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HEPARANASE ACTIVITY  
 ; FILE REFERENCE: 01/22335  
 ; CURRENT APPLICATION NUMBER: US/09/930,218  
 ; CURRENT FILING DATE: 2001-08-16  
 ; PRIOR APPLICATION NUMBER: 09/666,390  
 ; PRIOR FILING DATE: 2000-09-20  
 ; NUMBER OF SEQ ID NOS: 16  
 ; SOFTWARE: Patent In version 3.1  
 ; SEQ ID NO 2  
 ; LENGTH: 536  
 ; TYPE: PRT  
 ; ORGANISM: Rattus rattus  
 ; US-09-930-218-2

Query Match 37.7%; Score 1165; DB 9; Length 536;  
 Best Local Similarity 45.1%; Pred. No. 7,56-106; Indels 28; Gaps 7;  
 Matches 248; Conservative 87; Mismatches 187;

QY 42 GDRRLPVDRAGLKEKTLILLDVSTKNPVRYVNEFLSLQDPSIHDGMLDPLSSKRL 99  
 DB 12 GDRRLPVDRAGLKEKTLILLDVSTKNPVRYVNEFLSLQDPSIHDGMLDPLSSKRL 71  
 QY 100 RLVTLARGLSPAFLRFGKRTDPLQFONLRNPAKSGGPGDYLYKNEEDIVRSDVALDKQ 159  
 DB 72 RLVTLARGLSPAFLRFGKRTDPLQFONLRNPAKSGGPGDYLYKNEEDIVRSDVALDKQ 123  
 QY 160 KCKGKIAHPVMTVLOREKAAQWHLVILKEQFNTYSNLTITARSIDKLVPADCSGLHL 219  
 DB 124 ----- RVSADVLAKLQWMPFQ-ELLILRQYREFKNSYSSSDMLYSFASKRL 175  
 QY 220 HIIFALNALRPNPNSWSSSALSILKYSASKYNIWELGNEPNRYTMHGRAVNGSOLG 279  
 DB 176 HIIFALNALRPNPNSWSSSALSILKYSASKYNIWELGNEPNRYTMHGRAVNGSOLG 235  
 QY 280 GRDYIOLKSLQPIRIRYSASLYGNIGRPRKNVIALLDGFMKVASIVDAVWQHCYIDGR 339  
 DB 236 GRDYIOLKSLQPIRIRYSASLYGNIGRPRKNVIALLDGFMKVASIVDAVWQHCYIDGR 294  
 QY 340 GRVYKMDLKTLLDLSDOIIRKIOKVNTYTPGKKILLEGVTTISAGTNNLSYSYAGF 399  
 DB 295 GRVYKMDLKTLLDLSDOIIRKIOKVNTYTPGKKILLEGVTTISAGTNNLSYSYAGF 354  
 QY 400 GFLMNTLGMLANOGIDVYIRHSFDFHGYNHLVDQNFPLPDYWLSTLYKRLIGPKYLAHV 459  
 DB 355 GFLMNTLGMLANOGIDVYIRHSFDFHGYNHLVDQNFPLPDYWLSTLYKRLIGPKYLAHV 414  
 QY 460 HVAQLOKRPGRVITDKLRIYACHTNNHNNYVRSITLFIINLHRSKKIKLAGTLDKLV 519  
 DB 415 HVAQLOKRPGRVITDKLRIYACHTNNHNNYVRSITLFIINLHRSKKIKLAGTLDKLV 465  
 QY 520 KLVHGYILOPYGEGLSKSVQUNGOPLVMDGTLPELKRPLRAGTLYIPVTMGFVYV 579  
 DB 466 KLVHGYILOPYGEGLSKSVQUNGOPLVMDGTLPELKRPLRAGTLYIPVTMGFVYV 525  
 QY 580 KVNNALACRYR 589  
 DB 526 KVNNALACRYR 535

```

RESULT 10
US-10-431-438-2
; Sequence 2, Application US/10431438
; Publication No. US20030180788A1
; GENERAL INFORMATION:
; APPLICANT: goldshmidt, orit
; APPLICANT: pecker, iris
; APPLICANT: vlodavsky, israel
; APPLICANT: israel, michael
; TITLE OF INVENTION: AVIAN AND REPTILE DERIVED POLYNUCLEOTIDE ENCODING A POLYPEPTIDE H
; FILE REFERENCE: 26013
; CURRENT APPLICATION NUMBER: US/10/431,438
; CURRENT FILING DATE: 2003-05-08
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Rattus rattus
US-10-431-438-2

Query Match      37.7%; Score 1165; DB 14; Length 536;
Best Local Similarity 45.1%; Pred. No. 7.5e-106;
Matches 248; Conservative 87; Mismatches 187; Indels 28; Gaps 7;

QY 42 GDRBPLPYDRAAG-LKEKTLILDVSTKNPVYTVNENFLSLQDPSIIHD-GWIDFLSSK 99
DB 12 GLRALTLGTGTAAGTAFTDGVVDLEFYTKRLFGQSVSPFLSTITIDASLATDPRFLTFGSP 71
QY 100 RLVTIARSLSPAFIRFGKRTDFOFQNLKRNPAKSGPGEDYLYKNYEDDIYRSVDVALD 159
DB 72 RLRLARSLSPAFIRFGKRTDFOFQNLKRNPAKSGPGEDYLYKNYEDDIYRSVDVALD 123
QY 160 KQKQCKIAQHPDVMVLQREKRAQWHLVILKEQFSNTYSNLIITARSIDKLYNFDSCGL 219
DB 124 -----RVSADVLRKIQWEMPFG-ELLILREQYQREPKNSTYRSASVDMLYSFAKCSRL 175
QY 220 HILPALNALRPNNSNNSSSALSLIKTSASKNTNISWELGNEPNNTYRTHGKRAVNSQL 279
DB 176 DLIFLNLALRTPDLRNNSSNAQOLLNCCSKGNISWELGNEPNNSFWKKAQISIDQL 235
QY 280 GKDYIOLKSLIOPRIYRSASLYGPNIGRPKNVIALLDGFMKAGSTVDVAVTWOHCYID 339
DB 236 GEDPEYELKHLQK-SAFQNAKLYGPDIGQPKGVKILRSFLKAGGEVIDSLTWHHYLN 294
QY 340 GRVAVKNDFLKRLDITLSDOIIRKIQKVNTYTPPEKTIWEGVVTTSAGGNNLSDSYAA 399
DB 295 GRVAVKEDFLSSDVLDITFLISQKILKTKEMTPPKKVMLETSASVGGGAPLLSNTFAA 354
QY 400 GELMNTLGLMLANOGIDIVIRHSPFDHGYNLVDONFNP.LPDYWL.SILYKELIGPKYLA 459
DB 335 GFMMLDKIGLSAQOLGIEVVMKQVFFGAGNYHLDENFEP.LPDYWL.SILYKELIGPKYLA 414
QY 460 HVAQLQRRKPRGRVTRDKLRIYAHCTNNHNNYVRSITFLINLHRSRKIKLAGTLRD 519
DB 415 RYKQPD-----RSKLRVYLHCTNYHPRYRGD.LTYLVNLHNTYKHLKPPMFS 465
QY 520 KLVHGYLQPYGQGLKSKSVOLNGOPLVMVDGTLPELKRPRLRAGRTLVIPVTMGFY 579
DB 466 RFVDKTLKPFSGSDGLSKSVOLNGOPLVMVDGTLPELKRPRLRAGRTLVIPVTMGFY 525
QY 580 VVKVNVNALAC 589
DB 526 VTRNAKIAAC 535

```

```

; APPLICANT: pecker, iris
; APPLICANT: vlodavsky, israel
; APPLICANT: israel, michael
; TITLE OF INVENTION: AVIAN AND REPTILE DERIVED POLYNUCLEOTIDE ENCODING A POLYPEPTIDE H
; FILE REFERENCE: 01/2235
; CURRENT APPLICATION NUMBER: US/09/930,218
; CURRENT FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 09/666,390
; PRIOR FILING DATE: 2000-09-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-930-218-1

Query Match      37.7%; Score 1163.5; DB 9; Length 535;
Best Local Similarity 46.2%; Pred. No. 1.1e-105;
Matches 244; Conservative 80; Mismatches 177; Indels 27; Gaps 6;

QY 63 LDVSTKNPVYTVNENFLSLQDPSIIHD-GWIDFLSSKRLVTIARGLSPAFIRFGKRTD 121
DB 33 LEFYTKRPLRSVSPFLSTITIDASLATDPRFLTFGSPRLALARGSPAYIRFGGTVD 92
QY 122 FLQFQNLKRNPAKSGPGEDYLYKNYEDDIYRSVDVALDKQKCKIAQHPDVMVLQREKA 181
DB 93 FLIF-----DPDEPTSEERSYKWSQVNDICRSEV-----SAAVLRKIQWEMP 137
QY 182 AQMHVILKEQFSNTYSNLIITARSIDKLYNFDSCGHLIFALNALRPNNSNNSSSA 241
DB 138 FG-ELLILREQYQREPKNSTYRSASVDMLYSFAKSGGIDILFGNALRLTDLRNNSSNA 196
QY 242 LSLIKTSASKNTNISWELGNEPNNTYRTHGKRAVNSQLGKYIOLKSLIOPRIYRSASL 301
DB 197 QLLIYCCSKSGYNISWELGNEPNNSFWKKAHLLIDQLGDEDFVEHLKILQK-SAFQNAK 255
QY 302 YGPNIGRPKNVIALLDGFMKAGSTVDVAVTWOHCYIGRVAVKWDPLFKRLDITLSDOI 361
DB 256 YGPDIGQPKGVKILRSFLKAGGEVIDSLTWHHYLNKRLATKDFLSSDMLDITLSV 315
QY 362 RKIQKVVNTYTPPEKTIWEGVVTTSAGGNNLSDSYAAQFMMLNTLGLMLANOGIDIVIRH 421
DB 316 OKILKVTKEITQKKVMLETSASVGGGAPLLSNTFAAGFMMLDKIGLSAQWIEVVMKQ 375
QY 422 SPFDHGYNLVDONFNP.LPDYWL.SILYKELIGPKYLAHVAGLQRRKPRGRVTRDKLRIY 481
DB 376 VFFGAGNYHLDENFEP.LPDYWL.SILYKELIGPKYLAHVAGLQRRKPRGRVTRDKLRIY 426
QY 482 AHCTNNHNNYVRSITFLINLHRSRKIKLAGTLRDKLVHGYLQPYGQGLKSKSVQ 541
DB 427 LHCTNYHPRYRGD.LTYLVNLHNTYKHLKPPFLFRSPVDTYLLKSGPGGLSKSVQ 486
QY 542 LMGQPLVMVDGTLPELKRPRLRAGRTLVIPVTMGFYVVKVNVNALAC 589
DB 487 LMGQPLVMVDGTLPELKRPRLRAGRTLVIPVTMGFYVVKVNVNALAC 534

```

```

RESULT 11
US-09-930-218-1
; Sequence 1, Application US/09930218
; Patent No. US20020034810A1
; GENERAL INFORMATION:
; APPLICANT: goldshmidt, orit

```

; PRIOR FILING DATE: 1997-09-02  
 ; PRIOR APPLICATION NUMBER: US 09/109,386  
 ; PRIOR FILING DATE: 1998-07-10  
 ; PRIOR APPLICATION NUMBER: PCT/US98/17954  
 ; PRIOR FILING DATE: 1998-08-31  
 ; NUMBER OF SEQ ID NOS: 47  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 44  
 ; LENGTH: 535  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-776-874A-44

Query Match 37.7%; Score 1163.5; DB 9; Length 535;  
 Best Local Similarity 46.2%; Pred. No. 1,1e-105;  
 Matches 244; Conservative 80; Mismatches 177; Indels 27; Gaps 6;

QY 63 LDVSTKPVRTVNEFLSLQDPSIIHD-GWLDPLSKRLVTLARGISPAFLRFGKRTD 121  
 DB 33 LEFTKRLRSVSPFLSITIDASLATDPRFLTGLSPRLALARGISPAFLRFGKRTD 92  
 QY 122 FLOFONLNPASRGSGGPPDYLLKNYEDDIYRSDVALDKQCKIAGHPVMTLQREKA 181  
 DB 93 FLIF----DPDKPTSEERSYKSNVNDICRSEPV-----SAAVLRKLQVEMP 137  
 QY 182 AQMHVILKEQPSNTYSNLIITARSIDKLYNFADCGSLHIFALNALRPNNSMNSSA 241  
 DB 138 FQ-ELLRLREYQKEFNKSTYSRSSVDMLYSPACSGDLIFGALNLTPTDLRMSSNA 196  
 QY 242 LSLKYSASKKKNISWELGNEPNRYTMGRAVNGSQKDYIOLKSLQPIRYSASL 301  
 DB 197 QLLDLYCSSKKNISWELGNEPNRSFWKKAHLIDGLQGEDFVEIHLKLR-SAFONAKL 255  
 QY 302 YGPNIGRPKNVITALLDFPMKAGSTVDVAVTWQHCYIDGRVYKMDLKTLLDTSQI 361  
 DB 256 YGPDIGQPRGKTVKLLRSEFLKAGGEVIDSLTMHHYILNGRATYEDFLSSDALDTFLISV 315  
 QY 362 RKIOKVNTYTPGKKIMLEGVVTTSAGGTNNLSDSYAGFLMNTLGMLANOGIDVYIRH 421  
 DB 316 QKILKTKETITPGKKVMTGETSSAYGGAPLSTNTPAAGFWMIDKLGISAQMGIEVVMRQ 375  
 QY 422 SFPDHGVNHLVDONNPLPDYMLSLYKRLIGPKVLAHVAGLQKRPGRVIRDKLRIY 481  
 DB 376 VFFGAGNYHLVDENEPPLDYMLSLFKVLGPRVLSRVKGPD-----RSKLRVY 426  
 QY 482 AHCNHNHNHYRGSITLFIINLHRSRKKIKLAGTLRDKLVHQYLLQPYGQGLKSKSVQ 541  
 DB 427 LHCTNVHPRYQEGDLTLVNLHNTGHLKVPPLFRKPPVDYLLKPSGPDGLSKSVQ 486  
 QY 542 LINGQPLVWVDGTLPELKRPLRAGRTVIPPVTMGFVVVKNVNALAC 589  
 DB 487 LINGQILKNVDEQTLPALTEKPLPAGSALSLPAFSYGFVIRNAKTIAC 534

RESULT 13  
 US-09-988-113-44  
 ; Sequence 44, Application US/09988113  
 ; Patent No. US20020168749A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pecker, Itis  
 ; APPLICANT: Vlodavsky, Israel  
 ; APPLICANT: Feinstein, Elena  
 ; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY  
 ; TITLE OF INVENTION: EXPRESSION OF SAME IN GENETICALLY MODIFIED CELLS  
 ; FILE REFERENCE: 01/22781  
 ; CURRENT APPLICATION NUMBER: US/09/988,113  
 ; PRIOR FILING DATE: 2001-11-19  
 ; PRIOR APPLICATION NUMBER: US 09/776,874  
 ; PRIOR FILING DATE: 2001-02-06  
 ; PRIOR APPLICATION NUMBER: US09/258,892  
 ; PRIOR FILING DATE: 1999-03-01  
 ; PRIOR APPLICATION NUMBER: PCT/US98/17954  
 ; PRIOR FILING DATE: 1998-08-31

; PRIOR APPLICATION NUMBER: US 09/109,386  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: US 08/922,170  
 ; PRIOR FILING DATE: 1997-09-02  
 ; NUMBER OF SEQ ID NOS: 47  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 44  
 ; LENGTH: 535  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 ; US-09-988-113-44

Query Match 37.7%; Score 1163.5; DB 9; Length 535;  
 Best Local Similarity 46.2%; Pred. No. 1,1e-105;  
 Matches 244; Conservative 80; Mismatches 177; Indels 27; Gaps 6;

QY 63 LDVSTKPVRTVNEFLSLQDPSIIHD-GWLDPLSKRLVTLARGISPAFLRFGKRTD 121  
 DB 33 LEFTKRLRSVSPFLSITIDASLATDPRFLTGLSPRLALARGISPAFLRFGKRTD 92  
 QY 122 FLOFONLNPASRGSGGPPDYLLKNYEDDIYRSDVALDKQCKIAGHPVMTLQREKA 181  
 DB 93 FLIF----DPDKPTSEERSYKSNVNDICRSEPV-----SAAVLRKLQVEMP 137  
 QY 182 AQMHVILKEQPSNTYSNLIITARSIDKLYNFADCGSLHIFALNALRPNNSMNSSA 241  
 DB 138 FQ-ELLRLREYQKEFNKSTYSRSSVDMLYSPACSGDLIFGALNLTPTDLRMSSNA 196  
 QY 242 LSLKYSASKKKNISWELGNEPNRYTMGRAVNGSQKDYIOLKSLQPIRYSASL 301  
 DB 197 QLLDLYCSSKKNISWELGNEPNRSFWKKAHLIDGLQGEDFVEIHLKLR-SAFONAKL 255  
 QY 302 YGPNIGRPKNVITALLDFPMKAGSTVDVAVTWQHCYIDGRVYKMDLKTLLDTSQI 361  
 DB 256 YGPDIGQPRGKTVKLLRSEFLKAGGEVIDSLTMHHYILNGRATYEDFLSSDALDTFLISV 315  
 QY 362 RKIOKVNTYTPGKKIMLEGVVTTSAGGTNNLSDSYAGFLMNTLGMLANOGIDVYIRH 421  
 DB 316 QKILKTKETITPGKKVMTGETSSAYGGAPLSTNTPAAGFWMIDKLGISAQMGIEVVMRQ 375  
 QY 422 SFPDHGVNHLVDONNPLPDYMLSLYKRLIGPKVLAHVAGLQKRPGRVIRDKLRIY 481  
 DB 376 VFFGAGNYHLVDENEPPLDYMLSLFKVLGPRVLSRVKGPD-----RSKLRVY 426  
 QY 482 AHCNHNHNHYRGSITLFIINLHRSRKKIKLAGTLRDKLVHQYLLQPYGQGLKSKSVQ 541  
 DB 427 LHCTNVHPRYQEGDLTLVNLHNTGHLKVPPLFRKPPVDYLLKPSGPDGLSKSVQ 486  
 QY 542 LINGQPLVWVDGTLPELKRPLRAGRTVIPPVTMGFVVVKNVNALAC 589  
 DB 487 LINGQILKNVDEQTLPALTEKPLPAGSALSLPAFSYGFVIRNAKTIAC 534

RESULT 14  
 US-10-341-582-44  
 ; Sequence 44, Application US/10341582  
 ; Publication No. US20030161823A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Neta Ilan  
 ; APPLICANT: Israel Vlodavsky  
 ; APPLICANT: Oron Jacoby-Zeevi  
 ; APPLICANT: Itis Pecker  
 ; TITLE OF INVENTION: THERAPEUTIC AND COSMETIC USES OF HEPARANASES  
 ; FILE REFERENCE: 25449  
 ; CURRENT APPLICATION NUMBER: US/10/341,582  
 ; PRIOR FILING DATE: 2003-01-14  
 ; NUMBER OF SEQ ID NOS: 47  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 44  
 ; LENGTH: 535  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 ; US-10-341-582-44

Query Match 37.7%; Score 1163.5; DB 14; Length 535;  
 Best Local Similarity 46.2%; Pred. No. 11e-105;  
 Matches 244; Conservative 80; Mismatches 177; Indels 27; Gaps 6;

QY 63 LDVSTKPVATVNFENFLSLQDPSIHD-GWLDPLSKRLVTLARGLSPAFRFGSKRD 121  
 DB 33 LEFTYKRLPSVSPFSITITDASLATDPRFLTFLGSPRLARGLSPAYLRFQGTKD 92  
 QY 122 FLOFQNLRNPAKSRGGPDPYLLKNYEDDIVRSQVADLDKQCKIAQHPDVMVLQREKA 181  
 DB 93 FLIF-----DDKEPTSEERSYKMSQVNHDCRSEPV-----SAAVLKRLQVEMP 137  
 QY 182 AQMHVLVLEKQFSTYTNLLITARSLDKYNPADCGSLHLIPALNLRPNNSWNSSA 241  
 DB 138 FO-ELLILREQYQKFNSTYSRNSVDMYSPACSGDLDFGLNALLRFPDLRMSNA 196  
 QY 242 LSLKYSASKKYNISWEIGNEPNYRTMGRVAVNGSOLGQDYIQLKSLQPIRISPA 301  
 DB 197 QLLDYGSSKYNISWEIGNEPNRFSWKKAHLLIDGLQGEDFVELHKLQR-SAFQNAKL 255  
 QY 302 YGPNIGRPRKAVIALLDGFMKAVASTVDAYTWQHCYIDGRVVKMDFLKTLLDLSQ 361  
 DB 256 YGPDIGQPRGKTVKLLRSFLKAGGEVIDSLTWHHYINGRIATKEDFLSDALDTFLSV 315  
 QY 362 RKIQKVNNTYTPGKKIMLEGVVTTSAGTNNLSDSYAAGFLMNTLGMLANOGIDVIRH 421  
 DB 316 QKILKVTKEITPGKKVWLGERTSSAYGGAFLSVTPAAGFWMIDLKGLSOMGIEVVRQ 375  
 QY 422 SFPHGYNHLYDQFNPLPDYWSLLYKRLIGPKYLAHVAGLQKRRPRGVIRDKLRIY 481  
 DB 376 VFFGAGNYHLVDENEPFLPDYWSLLFKQVGPVLLSRVGPD-----RSKLRYV 426  
 QY 482 ACHTNHNNYVRGSITFLFINLHRSRKKIAGTLRDKLVHOYLLOPYQOEGILSKSVQ 541  
 DB 427 LHCTNVHPRQOEGDLTYVNLHNVTGHLKVPPLPRKPDYTLILKPSGPDGLSKSVQ 486  
 QY 542 LINGPLVWVDDGTLPELKRPLRAGRTLVIPVTMGFYVKNVNALAC 589  
 DB 487 LINGILKMWDEQTLPALTEKPLPAGSALSLPASFYGFVIRNAXIAAC 534

RESULT 15  
 US-10-384-451-44  
 ; Sequence 44, Application US/10384451  
 ; Publication No. US20030170860A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pecker, Irls  
 ; APPLICANT: Vlodavsky, Israel  
 ; APPLICANT: Feinstein, Elena  
 ; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY  
 ; FILE REFERENCE: 25718  
 ; CURRENT APPLICATION NUMBER: US/10/384,451  
 ; CURRENT FILING DATE: 2003-03-10  
 ; NUMBER OF SEQ ID NOS: 47  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO: 44  
 ; LENGTH: 535  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 US-10-384-451-44

Query Match 37.7%; Score 1163.5; DB 14; Length 535;  
 Best Local Similarity 46.2%; Pred. No. 11e-105;  
 Matches 244; Conservative 80; Mismatches 177; Indels 27; Gaps 6;

QY 63 LDVSTKPVATVNFENFLSLQDPSIHD-GWLDPLSKRLVTLARGLSPAFRFGSKRD 121  
 DB 33 LEFTYKRLPSVSPFSITITDASLATDPRFLTFLGSPRLARGLSPAYLRFQGTKD 92  
 QY 122 FLOFQNLRNPAKSRGGPDPYLLKNYEDDIVRSQVADLDKQCKIAQHPDVMVLQREKA 181

DB 93 FLIF-----DDKEPTSEERSYKMSQVNHDCRSEPV-----SAAVLKRLQVEMP 137  
 QY 182 AQMHVLVLEKQFSTYTNLLITARSLDKYNPADCGSLHLIPALNLRPNNSWNSSA 241  
 DB 138 FO-ELLILREQYQKFNSTYSRNSVDMYSPACSGDLDFGLNALLRFPDLRMSNA 196  
 QY 242 LSLKYSASKKYNISWEIGNEPNYRTMGRVAVNGSOLGQDYIQLKSLQPIRISPA 301  
 DB 197 QLLDYGSSKYNISWEIGNEPNRFSWKKAHLLIDGLQGEDFVELHKLQR-SAFQNAKL 255  
 QY 302 YGPNIGRPRKAVIALLDGFMKAVASTVDAYTWQHCYIDGRVVKMDFLKTLLDLSQ 361  
 DB 256 YGPDIGQPRGKTVKLLRSFLKAGGEVIDSLTWHHYINGRIATKEDFLSDALDTFLSV 315  
 QY 362 RKIQKVNNTYTPGKKIMLEGVVTTSAGTNNLSDSYAAGFLMNTLGMLANOGIDVIRH 421  
 DB 316 QKILKVTKEITPGKKVWLGERTSSAYGGAFLSVTPAAGFWMIDLKGLSOMGIEVVRQ 375  
 QY 422 SFPHGYNHLYDQFNPLPDYWSLLYKRLIGPKYLAHVAGLQKRRPRGVIRDKLRIY 481  
 DB 376 VFFGAGNYHLVDENEPFLPDYWSLLFKQVGPVLLSRVGPD-----RSKLRYV 426  
 QY 482 ACHTNHNNYVRGSITFLFINLHRSRKKIAGTLRDKLVHOYLLOPYQOEGILSKSVQ 541  
 DB 427 LHCTNVHPRQOEGDLTYVNLHNVTGHLKVPPLPRKPDYTLILKPSGPDGLSKSVQ 486  
 QY 542 LINGPLVWVDDGTLPELKRPLRAGRTLVIPVTMGFYVKNVNALAC 589  
 DB 487 LINGILKMWDEQTLPALTEKPLPAGSALSLPASFYGFVIRNAXIAAC 534

Search completed: May 6, 2004, 13:51:38  
 Job time : 51 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

# OM protein - protein search, using sw model

Run on: May 6, 2004, 13:42:11 ; Search time 20 Seconds  
(without alignments)  
2847.269 Million cell updates/sec

Title: US-10-088-676-2

Perfect score: 3088  
Sequence: 1 MRYLCAPFEMPSSNSRPPA.....PYTMGFYVKNVNALACRYR 592

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:\*

1: PIR1:1  
2: PIR2:1  
3: PIR3:1  
4: PIR4:1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2446	79.2	480	2 JCT506	heparanase protein
2	382	12.4	521	2 T45608	hypothetical prote
3	155	5.0	190	2 T01953	hypothetical prote
4	146	4.7	1260	2 T14022	reverse transcript
5	128	4.1	1027	2 AC1841	glycerophosphoryl
6	123	4.0	2026	1 OYBY	adenylate cyclase
7	115	3.7	923	2 S50462	hypothetical prote
8	114.5	3.7	837	1 A31842	endo-1,4-beta-xyla
9	106.5	3.4	1839	1 OYBY	adenylate cyclase
10	106	3.4	484	2 C88264	protein kin-15 (lm
11	106	3.4	488	2 T14430	hypothetical prote
12	105	3.4	747	2 S66959	hypothetical prote
13	105	3.4	835	2 C97322	probable alpha-ara
14	104.5	3.4	839	2 S73548	MG422 homolog C12
15	104.5	3.4	1280	2 B34087	hypothetical prote
16	104	3.4	635	2 T31679	hypothetical prote
17	101	3.3	535	2 S58740	heparanase protein
18	101	3.3	779	2 B42375	endopeptidase la
19	101	3.3	2244	2 F90563	hypothetical prote
20	101	3.3	2971	2 T08026	hypothetical prote
21	100.5	3.3	498	2 F83329	probable outer mem
22	100.5	3.3	566	2 T33042	hypothetical prote
23	100.5	3.3	630	2 C71374	probable glucosyl
24	100.5	3.3	878	2 A55201	meiosis-specific p
25	100.5	3.3	1259	4 S44887	retrovirus-related
26	100.5	3.3	3343	2 S44887	ZK112.7 protein -
27	99	3.2	785	2 F97001	endoglucanase fami
28	99	3.2	785	2 B72608	probable hyua APE1
29	99	3.2	871	2 S68482	probable membrane

## ALIGNMENTS

### RESULT 1

JCT506

heparanase protein 2a - human

C/Date: 17-Nov-2000 #sequence\_revision 17-Nov-2000 #text\_change 01-Dec-2000

C/Accession: JCT506

R/McKenzie, E.; Tyson, K.; Stamps, A.; Smith, P.; Turner, P.; Barry, R.; Hircock, M.; Pat

Biochem. Biophys. Res. Commun. 276, 1170-1177, 2000

A/Title: Cloning and expression profiling of Hpa2, a novel mammalian heparanase family m

A/Reference number: JCT506

A/Accession: JCT506

A/Molecule type: mRNA

A/Residues: 1-480 <MCK>

A/Cross-references: GB:AF282885

A/Comment: This protein, a intracellular membrane-bound enzyme, has biological and therai

C/Genetics:

A/Gene: hpa2a

A/Map position: 10q23-10q24

C/Keywords: heparin binding; membrane bound

Query Match 79.2% Score 2446; DB 2; Length 480;

Best Local Similarity 80.9%; Pred. No. 1.5e-183;

Matches 479; Conservative 1; Mismatches 0; Indels 112; Gaps 1;

QY	1	MRYLCAPFEMPSSNSRPPACIALGALYLAIIHISSSQAGDRRLPVDAAGIKETL	60
DB	1	MRYLCAPFEMPSSNSRPPACIALGALYLAIIHISSSQAGDRRLPVDAAGIKETL	60
QY	61	ILDDVSTKNPVATVVENFLSLQDPSIHDGWLDFSSKRLVTLARGSPAFLEFGKRT	120
DB	61	ILDDVSTKNPVATVVENFLSLQDPSIHDGWLDFSSKRLVTLARGSPAFLEFGKRT	120
QY	121	DLQFQNTNRPKASGSGGPGDYLLKNYEDDIYRSVVALDKQCKIAQHPDMLVLOREK	180
DB	121	DLQFQNTNRPKASGSGGPGDYLLKNYEDDIYRSVVALDKQCKIAQHPDMLVLOREK	180
QY	181	AAQWMLVLKQFSUTYSNLILFARSLKLVNFADCSGLHIFALNALRPNNSWNSS	240
DB	181	AAQWMLVLKQFSUTYSNLILFARSLKLVNFADCSGLHIFALNALRPNNSWNSS	240
QY	241	ALSLKYSASKKYNISWEIENPNRYTMHGRAVNGSLQKQDYQLKSLQPIRIYSPAS	300
DB	241	ALSLKYSASKKYNISWEIENPNRYTMHGRAVNGSLQKQDYQLKSLQPIRIYSPAS	300
QY	301	LYGPNIGRPKNVIALDGMKVAAGSTYDAVWQHCYIDGRVVKYMDLKTRLDPTLSDQ	360
DB	301	LYGPNIGRPKNVIALDGMKVAAGSTYDAVWQHCYIDGRVVKYMDLKTRLDPTLSDQ	360
QY	361	IRIKQKVVNTYTPGKKIMLEGVVTSAGGTNNLSDSYAAGFLMNTLQMLANOGIDVIR	420
DB	361	IRIKQKVVNTYTPGKKIMLEGVVTSAGGTNNLSDSYAAGFLMNTLQMLANOGIDVIR	420
QY	421	IRIKQKVVNTYTPGKKIMLEGVVTSAGGTNNLSDSYAAGFLMNTLQMLANOGIDVIR	480
DB	421	IRIKQKVVNTYTPGKKIMLEGVVTSAGGTNNLSDSYAAGFLMNTLQMLANOGIDVIR	480

```

QY 421 HSPFDHGNHLYVDQNFPELPDYMLSLKYRLIGPKYLAHVAGLQKRPGRVIRDKLR 480
DB 309 HSPFDHGNHLYVDQNFPELPDYMLSLKYRLIGPKYLAHVAGLQKRPGRVIRDKLR 368
QY 481 YAHCTNHNHNHYVGSITLFTINLHRSRKKIKLAGTLRDLVHQVYLQPYGQGLSKSV 540
DB 369 YAHCTNHNHNHYVGSITLFTINLHRSRKKIKLAGTLRDLVHQVYLQPYGQGLSKSV 428
QY 541 QLNQGPVWVDGTLPELKPRLPAGRTVIRPTWFWYKVNVLACGR 592
DB 429 QLNQGPVWVDGTLPELKPRLPAGRTVIRPTWFWYKVNVLACGR 480

RESULT 2
T45608
hypothetical protein F13G24.30 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T45608
R:Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voel, M.; Robben, J.; Volckaert, G.; Ba
submitted to the Protein Sequence Database, December 1999
A:Reference number: 223009
A:Accession: T45608
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-521 <BEV>
A:Cross-references: EMBL:AL133421
A:Experimental source: cultivar Columbia; BAC clone F13G24
C:Genetics:
A:Map position: 5
A:introns: 53/3; 66/1; 127/2; 177/1; 256/1; 319/2; 361/2; 394/3
A>Note: F13G24.30

Query Match 12.4%; Score 382; DB 2; Length 521;
Best Local Similarity 25.8%; Pred. No. 8.1e-22;
Matches 154; Conservative 76; Mismatches 187; Indels 180; Gaps 28;

QY 75 NENLSLQLD--PSIIHD-----GWLDFLS--SKRLVTLA-RGLSPALRGGGR 119
DB 22 DENFVCAITLDMWP--HDKCNVDQCPWGSVSSVIMDLTRPLTLTAIFRPLRIRGGSL 78
QY 120 TDFLQFO--NLRNPA-----KSRGSPDPYYLKNYEDDIVSDVALDKQCKIAQHPDV 173
DB 79 QDQYIYDVGNLKTCPRFQKNGS-----LFGSKXC----- 110
QY 174 LVLRERAAQMHVLLKQPSNTYSNLTIRASIDKLYNADSGGLHTALNMR-- 229
DB 111 -----LHM-----KRWDELNSFLTATGAVVTGMLALRGRK 142
QY 230 ---RNPNNMNSSSALSLKYSASKKTYNI--SWEIENPNPNYRTWNGRAVNGSOLGKDYIQ 285
DB 143 LRGRAMGAMWHNTQPFNAVTVSKGYVIDSWERGNLSGSV--GASVARELYGDLIV 200
QY 286 LKSLQIPRIYRSALSYGENTGRPRKVALLDGF-----MKVAG--STDVAVTMOH 335
DB 201 LKQVIN--KYKNSWLKRPILVAP-----GGFYEQQWYTKLEISGSSVDDVAV--H 249
QY 336 CYIIGR-----VYKWDFFLKTLLDLSQIRIKQKVVNTYPPGKKIMEGVVTSAG-- 388
DB 250 IYNGSGNDPLVAKIND--PSYLSQVSKTFKQVNTIQEHG--WASPMVSGSGA 302
QY 389 ---GTNNLSDSYAAGFLMANTLGMLANOGIDVIRHSFPDHGYNHLYVDQNFPLDYMLS 445
DB 303 YNSGRHVSDFTFISFWLDTQIGSARHNTKYVCRQTLVGGFYGLKKGFFVNPDPYISA 362
QY 446 LLYRRLIGPKYLAHVAGLQKRPGRVIRDKLRITVYAHCTNHNHNHYVGSITLFTINLH 505
DB 363 LLMHRLMGKYLAVOTG-----PFQLAVYAHCSG-----PAGVITLLILNLS 405
QY 506 -----RSRRKIKLAGTLRDLV-----HOYLIQPYGQ 533
DB 406 NQSDFTVSVSNGINVLNABSRKKSLDLTLKRPFSWIGSKASDGYINREYHLTP--EN 463

```

```

QY 534 G-LKSKSVQLNGOPLVWVDGTLPELKPRLPAGRTVIRPTWFWYKVNVLAC 589
DB 464 GVLRSKTMVNLGSKLKPATNDISLEPVRSVNPLNVLPLSSFLVLPFDSAC 520

RESULT 3
T01953
hypothetical protein T2L5.6 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 21-Jan-2000
C:Accession: T01953
R:Geisel, C.; Smith, A.; Le, T.
submitted to the EMBL Data Library, October 1998
A:Description: The sequence of A. thaliana T2L5.
A:Reference number: 214470
A:Accession: T01953
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-190 <GEI>
A:Cross-references: EMBL:AF096371; NID:g3695386; PID:g3695392
A:Experimental source: cultivar Columbia
C:Genetics:
A:Map position: 4
A:introns: 36/2; 69/3
A>Note: T2L5.6
C:Superfamily: Arabidopsis thaliana hypothetical protein T2L5.6

Query Match 5.0%; Score 155; DB 2; Length 190;
Best Local Similarity 24.9%; Pred. No. 0.0001;
Matches 49; Conservative 32; Mismatches 70; Indels 46; Gaps 6;

QY 420 HSPFDHGNHLYVDQNFPELPDYMLSLKYRLIGPKYLAHVAGLQKRPGRVIRDKLR 479
DB 12 RQSLIGNVGLNTTTFPNPDYVSALTWRQMRKALFTTSGTK-----KIR 60
QY 480 IYAHCTNHNHNHYVGSITLFTINLHRSR--KKIKLAGTLRDLVHQVYLQPY----- 530
DB 61 SYTHCARQSK-----GITVLMLNDTNTVAVAVELNNSF--SLRHTKHKYSKRASSQ 112
QY 531 ---GQEG-----LKSQSVQLNGOPLVWVDGTLPELKPRLPAGRTVIR 572
DB 113 LFGPNQYIQREBYHLTRKQDNLHSQTLNGLNALQVNSGDLPIEIRHINSTEPIITA 172
QY 573 PYTMGFYVAKVNALAC 589
DB 173 PYSLVFYHMRNVVPAC 189

RESULT 4
T14022
reverse transcriptase homolog - slime mold (Dictyostelium discoideum) retrotransposable
C:Species: Dictyostelium discoideum
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 11-May-2000
C:Accession: T14022
R:Winkler, T.; Tschepke, C.; de Hostos, E.L.; Jendretzke, A.; Dingermann, T.
Mol. Gen. Genet. 257, 655-661, 1998
A:Title: Tdd-3, a transfer RNA gene-associated poly(A) retrotransposon from Dictyostelium
A:Reference number: 217858; MUID:98265925; PMID:9604889
A:Accession: T14022
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1260 <MIN>
A:Cross-references: EMBL:AF002669; NID:g2558820; PID:g2558822; PIDN:AA048324.1

Query Match 4.7%; Score 146; DB 2; Length 1260;
Best Local Similarity 21.3%; Pred. No. 0.0099;
Matches 141; Conservative 101; Mismatches 229; Indels 192; Gaps 39;

QY 63 NPVRTVNEFSLQDPSIHDGWL--DFL--SKRLV-----TLARGSP 110
DB 94 NGIGILHNHNNQIKSPFIIEGRLLISDILIKOTTRILAIYAPQDPKRTIASLTLNK 153

```



```

QY 111 AFLFGGKRTDPLQFQNTL-RNPASRGSGPDYLYLKNYEDDIYRSVVALDKOKCKLIAOH 169
DB 154 HFNK-----QYHULTSNENKNDIILAGDFNCIDFNDNHTSND-----DQGNLTQS 199
QY 170 PDVVLVLR-----EKAQWHLVLLKEQFS-----NTYSNLI-LTARSLDLY-----NF 213
DB 200 PDEMAVTEAIRISNNIMDMDOQNKRPFTFRTINTNNLTFLRLERLDRIYANNSILNY 259
QY 214 -----ADCSGLHIFALNALARNNNW--NSSGALS--LLKXSASK 251
DB 260 SOLYRNLIIPKINDIPLSDFNFLTFTLANIQTNRR--WRLKSSILSSIMLK----- 313
QY 252 KYNISWEIAGNPNRYRTWGRVAVNGSQIGKDYIQLKSLQPIR-IYSRSLYSPNIGRPR 310
DB 314 --NIDFL-----NGY-----SREISSNENISFSQJNSLKNKIKQLYTEFQKQDYNKAN 363
QY 311 -KAVIALLD-GFMKVASSTYDAYTV-----QHC-----YIDGRVVK--VMDP 348
DB 364 IKULISLETFEYKQAFATLSAINESKKEBOLKQELANNCEETSLKYISARIKKRNDP 423
QY 349 LKTRLDLTLSDOJRIQKVVVYTY-----PGKIWLEGVTTSGAGTNLS 394
DB 424 TINAVKOTQGRITNKQILIEEYKYXSNLYDYKEDDPESHYELIEWVTVTRDSTWMDLE 483
QY 395 DSYAGFLW--LNTGLMLANOGIDVYTRHSEFDH-----GIN-----HLYDQNF 436
DB 484 NEFTSOEILEVYKQINLPHKSPGPDGIPNLFITHEKELAPILASAFNDTLRNPHLISKNY 543
QY 437 -----NLPDYWLSLYK--RLIGPKYLAHVHAGLQKRR-----PGR 472
DB 544 KEGIIITPKKQDELKRRPITLANCIYKISHKLNNEIIPILTYVINHNOXGAPR 603
QY 473 VIRDKLRVYACHTNHNHNVYSGITLFIINLHRSRKKIKLAGLRDKL--VHOYLLQPY 530
DB 604 FILNITISINELIYXNDKINGITITLYI-----SKDLTRSHVQSQITTHQHSQYI 658
QY 531 GQ-BGLSKS---VOLNGQPLWMDGTLP-ELKRLRLRGRVIVIPVYMGVYVKNV 555
DB 659 NLINMLTKSEARLEINGRT-----TIPEIK-RGVKQG-----DPLSPTLFLVIV-IE 704

QY 586 ALA 588
DB 705 ALA 707

RESULT 5
AC1841
glycerophosphoryl diester phosphodiesterase [imported] - Nostoc sp. (strain PCC 7120)
C/Species: Nostoc sp. PCC 7120
A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C/Accession: AC1841
R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A/Reference number: AB1807; MUID:21595285; PMID:11759840
A/Accession: AC1841
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1027 <CUR>
A/Cross-references: GB:BA000019; PIDN:BA077799.1; PID:G17135253; GSPDB:GN00179
A/Experimental source: strain PCC 7120
C/Genetics:
A/Gene: all0275

Query Match 4.1%; Score 128; DB 2; Length 1027;
Best Local Similarity 21.9%; Pred. No. 0.18;
Matches 127; Conservative 70; Mismatches 222; Indels 160; Gaps 30;

QY 62 LLDVSTKNPRTVNE-----NFLSQDPSII-HDGMWDLFSSKRLVTLAAGLSAPFL 113
DB 101 LSPDNMKIPQIYVNEASADRLTGADLDISFVIDXG-----SIWVGDFEFGYLL 151

```

```

QY 114 RFG--GKRTDFL-----QFQNLARNPASK-----RGSPG--PDYLYKNYE-----DIYRS 154
DB 154 HFDATGKLIDAPLAPPPBFTKTLIDGTAPREVITGRGASGRPHHTLESYCLALEQADFIER 211
QY 155 DVALDKOKGCKIAQHPDVLVYIOREKAQWHLVLLKEQFSNTYSNLI-LTARSLDLYN-- 212
DB 212 DLAVTKD--GVILARHEPALVALVNDGSV-----NFSVTTNNVQIAPFSRLKTVN 261
QY 213 -----FADCSGLHIFALNALARNP--NNSNSSASLSLKYSASKKNISWEIAGN 261
DB 262 LDGTEITGWFADFTLAIKELRALIERLFRDQSGNQFTPTLAEIIDLKQVAEATGK 321
QY 262 EPNRY-RTWGRVAV-----NGSQIGKDYIQLKSLQPIR-IYSRSLYSPNIGRPN 305
DB 322 KIGIYPERKHPYFPAQNTVYGTTEKINRNISQILIDLKANNFTDEPRITIGSEFVG-- 379
QY 306 IGRPRKNVIALIDGEMKVASSTV-----DAVWQHICYIDGRVVK--MDFLKTRLDLTL 357
DB 380 -----NKEIHDITMPNAGVDIPLVQLFDAIDVD--INGRLIETREYDFIVSGDTRTY 430
QY 356 SDQIRKIQV--VNTYTPGKKIMLEGVTTSGAGTNLSDSYAAGFLMLNTGLMLANOGI 415
DB 431 GD-LTPAGLAIAYADGIGPKRMIVSAGTDANN-----DQAD 471
QY 416 DVVIRHSEFDHGYNLVDQNRNPLPDYWLSLLYKRLIGPKYLA--VHYAGLQKRRPRGRV 473
DB 472 DV-----NKGDAVNDAD-----KTLLEPTLVQDAHVHAGLQVHP---YT 507
QY 474 IRDKLRVYACHTNHNH--NY-----VRSITLFIINLHRSRKKIKLAGT--LR--- 518
DB 508 FRDEERYLA--ANYGNELELYQQLFGVVALFDFPTIYDRVRDLSPGNIVASPO 565
QY 519 --DKLVHGYLLQPYGOEGLKSKSVOLNGQPLWMDGTLP 555
DB 566 NPDVLSGDAFNLGGSRFEGGATVASKTKLYMLEGTV 604

RESULT 6
OXYB
adenylate cyclase (EC 4.6.1.1) - yeast (Saccharomyces cerevisiae)
N/Alternate names: ATP pyrophosphate-lyase; protein J1401; protein YJL005W
C/Species: Saccharomyces cerevisiae
C/Date: 28-Dec-1987 #sequence_revision 08-Sep-1995 #text_change 21-Jan-2000
C/Accession: S56776; S56775; A24776; S05828; S55183
R/To Van, D.; Pereira, J.; Jacq, C.
submitted to the Protein Sequence Database, September 1995
A/Reference number: S56776
A/Accession: S56776
A/Molecule type: DNA
A/Residues: 1-1823 <DEH>
A/Cross-references: EMBL:Z49280; GSPDB:GN00010; MIPS:YJL005W
R/de Haan, W.; Grivell, L.A.; Smit, P.H.M.
submitted to the Protein Sequence Database, September 1995
A/Reference number: S56775
A/Accession: S56775
A/Molecule type: DNA
A/Residues: 673-2026 <ZAG>
A/Cross-references: EMBL:Z49280; GSPDB:GN00010; MIPS:YJL005W
R/Kataoka, T.; Broek, D.; Wiggler, M.
Cell 43, 493-505, 1985
A/Title: DNA sequence and characterization of the S. cerevisiae gene encoding adenylate
A/Reference number: A24776; MUID:86079531; PMID:2934138
A/Accession: A24776
A/Molecule type: DNA
A/Residues: 1261,'U', 263-547,'U', 549-591,'H', 593-708,'I', 710-961,'P', 963-1387,'S', 1389-
A/Cross-references: EMBL:M12057; NID:G171359; PIDN:AA14549.1; PID:G171360
A/Note: the authors translated the codon TTA for residue 262 as Ser, AGC for residue 311
2 as Leu, TCA for residue 1566 as Ala, GAG for residue 1659 as Ala, GGT for residue 1735
R/Asson, P.; Lenzen, G.; Jacquemin, J.M.; Danchin, A.
Curr. Genet. 10, 343-352, 1986
A/Title: Yeast adenylate cyclase catalytic domain is carboxy terminal.
A/Reference number: S05828; MUID:88165073; PMID:3327602

```

A/Accession: S05828  
 A/Molecule type: DNA  
 A/Residues: 1042-1426, 'D', 1428-1460, 'T', 1462-1955, 'V', 1957-2008, 'M', 2014-2015, 'TNEYKMLRT'  
 A/Cross-references: EMBL:X03449; NID:G3487; PIDN:CAA27175.1; PID:G3488  
 A/Submitter: M. Smits, P.H.M.; Grivell, L.A.  
 A/Reference number: 555183  
 A/Accession: S55183  
 A/Molecule type: DNA  
 A/Residues: 673-2026 <DEM>  
 A/Cross-references: EMBL:X87611; NID:G854567; PIDN:CAA60917.1; PID:G854568  
 A/Genetic: SGD:CYR1; CDC35; MIPS:YJL005W  
 A/Cross-references: SGD:S0003542; MIPS:YJL005W  
 A/Map position: 10L  
 C/Superfamily: yeast adenylate cyclase; leucine-rich alpha-2-glycoprotein repeat homolog  
 C/Keywords: CAMP biosynthesis; duplication; phosphorus-oxygen lyase; tandem repeat  
 F/669-1343/Region: leucine-rich 23-residue repeats  
 F/1065-1087/Domain: leucine-rich alpha-2-glycoprotein repeat homolog <LRR>  
 F/1610-2026/Domain: catalytic #status predicted <CAT>  
 F/1664-1749/Domain: yeast adenylate cyclase catalytic domain homolog <YACC>

Query Match 4.0%; Score 123; DB 1; Length 2026;  
 Best Local Similarity 22.6%; Pred. No. 1.3; Indels 162; Gaps 33;  
 Matches 141; Conservative 83; Mismatches 237;  
 37 ISSQAGRRPLPVDRAAGLKE-KTLLLDVSTKPVG--RTVNEFLSLQDPSIIHDGW 92  
 891 LINDQNELESLP---AGFVELKNLQDLDSNKFEMHPEVINYCNLQIDLSYKIGS 946  
 93 LDFLSSRLVTLAE-GISPAFLRFGKRTDF--LQFQNR-----NPAKSRGPG 140  
 947 LP-OSTYLYLAKMNSHKLNFIQDSEMTDLRTLNLRNRSIKTNASVQLMFLT 1005  
 141 DYVAKYEDIVRSADVLDKQCKIAQHPDVLVLRKAQW-HLVLKEQFSYTSN 139  
 1006 DNRISNEFD-----LPLKRALEIQENPITISFDFPKMTSLTLNKAQSSIPGE 1058  
 200 LILTARSLLDKLYNFADESGHLIFALNALRRPNNSNWSALSLKYSASKTYNSWEI 259  
 1059 LITLTSLEKLE-----LNQNNITRLPQEI-SKLTQVFLSVANNKLEYIPPEL 1106  
 260 GNEENNYRT--MHGRAVNGSLQKDYQLKSLQPIRYSRSL-----YGPNIQR 308  
 1107 -SOLSKFTLDLHNNIRDFPDGMENTELISLINSNAFGNSSENFYHMYGSLGSLK 1165  
 309 PRKNVIALDGFMKVAGSTVDVAVTWQ--HCYIDGRVVKVDFLTKRLDITLSQIRIKQ 366  
 1166 -----SLM--FLIADNQPDAMWPLNCFVN--LKVNLNLYNNFSDV---SHMKLES 1211  
 367 VVNTYTPQKIKWLEGVVTTTSAGTNNLSDSVAAAGFLW--LNTLGMLANQIDV-----V 418  
 1212 ITTELISGNKL-----TTLSGDT-----VLKNSSLKTLMLNSNOMLSLPAELSNL 1256  
 419 IRHSFFDHVNLVDONFNPLPDY--WLS--LTKRLIGP----- 454  
 1257 SOLSPVDGANQKNTSNHYDMWNRNKKELKTLNFSGNRFEIKFSISHDIDADLSL 1316  
 455 -----KTLAVHVAAGLQKRPKPRVIRDKLRIYAHCTNNHNN-----NLY----- 493  
 1317 TVLPQLKVLGLMDVTLNTTKVDPDENVFRLTFTTAIIIGMKRYGADTLGQRDIYSSSDVT 1376  
 494 -----RGSITLFIINLHRSRK-----IKLAGTLRK--LVHQYLLQPYGOE----- 533  
 1377 FERRRGNDDECLLCHHSKKNADYGNHISRIYVDYDKILIRQ--LERGDEDDNVIKT 1434  
 534 GLKSKSVOLNGQ--PLVMVDG 553  
 1435 ALRPSFLQINKEINGMNSVDNG 1457

RESULT 7  
 S50482

Hypothetical protein YER024W - Yeast (Saccharomyces cerevisiae)  
 C/Species: Saccharomyces cerevisiae  
 C/Date: 28-May-1993 #sequence\_revision 24-Feb-1995 #text\_change 19-Apr-2002  
 C/Accession: S50482  
 A/Submitter: R. Dietrich, F.S.  
 A/Reference number: 550433  
 A/Accession: S50482  
 A/Molecule type: DNA  
 A/Residues: 1-923 <DIE>  
 A/Cross-references: EMBL:U18778; NID:G603592; PID:G603616; GSPDB:GND0005; MIPS:YER024W  
 A/Genetic: SGD:YAT2; MIPS:YER024W  
 A/Cross-references: SGD:S0000826  
 A/Map position: 5R

Query Match 3.7%; Score 115; DB 2; Length 923;  
 Best Local Similarity 18.7%; Pred. No. 1.6;  
 Matches 115; Conservative 79; Mismatches 216; Indels 206; Gaps 27;  
 33 LHSLSQAGRRPLPVDRAAGLKEKTLLLDVSTKPVTVNENFSL--LQDPS--II 88  
 98 LHDINNQSTRB-----IQDDVLRPRPFLVADALPNTIQADRSAVLV 142  
 89 HDGWLDFLSSRLVTLARG--LSPALRFGKRTDFLOFQNL---RNPAKSRG----- 137  
 143 H-----SAARFISALQDILLPPDINATNGKPLSMAPFLNLFCTTSPVFORGEVENFD 195  
 138 -----DGPDY-----YLKNEYEDIVRSADVLDKQCKIAQHPDVLVQL 177  
 196 LNKPTAADDLEDPSSEDDNDEPQDPPDRKRKHEDLFTNGITLIRKHPSPSKHILI 255  
 178 REKAQWELVLKEQFSNTYNSLITARSLLDKLYN--PADSGHLIFALNAL----- 228  
 256 ISRGQYTLRYLV-----DSTWKIITYTAAELTTIRNHIKSSGIEKSTALGSLTSHSFRN 310  
 229 -----RNPNNNSNWSALSLKYSASKTYNSWEIENENNYRTMHGRAVNGSL 279  
 311 WKYARKLQKRYWELHIDSLFVLVDESOBERTN--DSDDTADISQMKNRKITERDK 368  
 280 GROYTQLSLQPIRYSRSLYGENIGRPRKNVIALDGFMKVAGSTVDA--VTMOC 336  
 369 KCTSANCK-----RVFYGTSLINSKGHGVSCSRWYDKQLVW--TADAKATVINDSF 420  
 337 YIDGRVVKVDFLTKRLDITLSQIRIKQKXVNTYTPQKIKV-----LEGV 383  
 421 TCDGSV--VLRFYS-----EITYESVLRARVNAAGDPQFSIMPNVTQMDPTKYLMTATI 474  
 384 TTSAGTNNLSDSVAAAGFLWNTLGMLANQIDVIRHSFFDHVNLVDONFNPLPDY 443  
 475 SADGGGSEIDPK-----LVNK-----IDWSPS----- 498  
 444 LSLKRLIGRPVLAHVHVAAGLQKRPKPRVIRDKLRIYAHCTNNHNNYVASTLFIIN 503  
 499 -----NINTHVLSETR-----LADLSKI-----DIYASITPL----- 528  
 504 LHSRKKIKLAGTLRDLKVHGYL-----LQPYGQGLK-SKSVOLNGQPLV 548  
 529 -----GRSAGRLGVKDSVVOVALQIAHYALYGMVFGLEPSTGPFKNRSRSFINIQOA 585  
 549 MYD-----DGT 554  
 586 LTELQFLISSIDGT 601

RESULT 8  
 A31842  
 endo-1,4-beta-xylosylase (EC 3.2.1.8) Z precursor - Clostridium thermocellum  
 N/Alternate names: xylosylase Z  
 C/Species: Clostridium thermocellum  
 C/Date: 31-Mar-1990 #sequence\_revision 11-Apr-1997 #text\_change 18-Jun-1999  
 C/Accession: A31842

R:Grepinet, O.; Chebrou, M.C.; Beguin, P.  
 J. Bacteriol. 170, 4582-4588, 1988  
 A:Title: Nucleotide sequence and deletion analysis of the xylanase gene (xynZ) of Clostridium  
 A:Reference number: A31842; MUID:8908072; PMID:3139632  
 A:Accession: A31842  
 A:Molecule type: DNA  
 A:Residues: 1-837 <GRE>  
 A:Cross-references: GB:M2624; NID:g144931; PIDN:AAA3286.1; PID:g144932  
 C:Gene: xynZ  
 C:Function:  
 A:Description: catalyzes the hydrolysis of 1,4-beta-xylosidic linkages in xylans  
 A:Pathway: xylan degradation  
 C:Superfamily: Clostridium endo-1,4-beta-xylanase Z; Clostridium cellulase repeat homolo  
 C:Keywords: duplication; extracellular protein; glycosidase; heat-stable protein; hydrol  
 F:128/Domain: signal sequence #status predicted <SIG>  
 F:29-837/Product: endo-1,4-beta-xylanase #status predicted <MAT>  
 F:326-419/Domain: Clostridium xylanase A repeat homology <CXA>  
 F:430-453/Domain: Clostridium cellulase repeat homology <CCR1>  
 F:464-487/Domain: Clostridium cellulase repeat homology <CCR2>  
 F:548-834/Domain: Streptomyces endo-1,4-beta-xylanase A homology <SXY>  
 F:645,754/Active site: Glu #status predicted

Query Match 3.7%; Score 114.5; DB 1; Length 837;  
 Best Local Similarity 21.0%; Pred. No. 1.5;  
 Matches 77; Conservative 54; Mismatches 128; Indels 107; Gaps 20;

QY 128 LRPAPKRG---GGPPDYLYKNEYEDIVASDVALDKQGGCKIAQHPDVMVLQREKAAQM 184  
 DB 518 LRDYAEKRGKIGTCVNYPPYNSDPYNS-----ILQREPSVVCNE----- 561  
 QY 185 HLVLKEQPSNTYSNLTITARSIDKLYNPADCSGL---HLIPALNALRNP---NSMW 236  
 DB 562 ---MKFDALQPRQNVDPFSGK--DQLLAFERNGMQRGHTLIWHN---QNPFWLITNGW 613  
 QY 237 NSSALSILKYSAS-----KKNISWELGNE-----PNNYRMRHRAVNSGLGKQYIQ 285  
 DB 614 NRSLAIVKQNHITVTWYHKGKIVEMDVANECMDSGNLSIMKNV---IGDDYLD 669  
 QY 286 LKSLQPIRIYSPAS-----LYG---PNIGRPKNVIALDGFMKVAGSTVDAVTWQH 335  
 DB 670 YAPR-----YARADDALLFNVDNIEDLG--PKSAVAPNMKSMKBERGVPIDGVGFC 722  
 QY 336 CYIDGRVVKWDFIKTLTLDLSDQIRKIQV-----VNTYTPKKITWLEGVYTTSA 387  
 DB 723 HFIING-----MSPEYLASIDONIKRYAIGVYSFTEIDIRIQS---ENPATAFG 770  
 QY 388 GGTNNLSDSYAAQFLMNTIGMLANGIDIVYIRHSFEDH-----GYNH--LVDPNF 436  
 DB 771 VQANNYKELMKI-----CLANPNONTFTVMWGFIDKXTWIPGTPPGYGNFLYDSNY 821  
 QY 437 NPLPDY 442  
 DB 822 NPKPAY 827

## RESULT 9

OYBYK  
 adenylate cyclase (EC 4.6.1.1) - yeast (Saccharomyces kluyveri)  
 N:Alternate names: adenyl cyclase  
 C/Species: Saccharomyces kluyveri  
 C/Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 21-Jan-2000  
 C/Accession: J01145; S14464  
 R:Young, D.; O'Neill, K.; Broek, D.; Wiggler, M.  
 Gene 102, 129-132, 1991  
 A:Title: The adenyl cyclase-encoding gene from Saccharomyces kluyveri.  
 A:Reference number: J01145; MUID:91323718; PMID:1864503  
 A:Accession: J01145  
 A:Molecule type: DNA  
 A:Residues: 1-1839 <YOU>  
 A:Cross-references: EMBL:X56042; NID:g4866; PIDN:CAA39513.1; PID:g4857  
 C:Genetics:  
 A:Gene: CYR1

C:Superfamily: yeast adenylate cyclase; leucine-rich alpha-2-glycoprotein repeat homology  
 C:Keywords: CAMP biosynthesis; duplication; phosphorus-oxygen lyase  
 F:1006-1029/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR>  
 F:1479-1564/Domain: yeast adenylate cyclase catalytic domain homology <YACC>

Query Match 3.4%; Score 106.5; DB 1; Length 1839;  
 Best Local Similarity 21.3%; Pred. No. 22;  
 Matches 132; Conservative 75; Mismatches 220; Indels 193; Gaps 33;

QY 56 KEKTLILLDVSTK---NPATVNE--NELQLDPSITIHGWDFLSKKLTARGLSP 110  
 DB 724 KLLQQLDLIDISSKRFVYPEVINSCTMLQIDISYNKHS--LP-VSINDVILAK----- 776  
 QY 111 APLRFGGKRT---DLPQONLR-----NPASRGSGP-----DYLYKNEYEDIVR 153  
 DB 777 -MNFNNRLTSVGLSGQMKRLRLNCRVTSIECHAPNOMLFTDNRIISFDDDLTR 835  
 QY 154 SDVALDKQGGCKIAQHPDVMVLQREKAAQMHVLYLKEQPSNTYSNLTITARSIDKLYNF 213  
 DB 836 -----LRTLELQNPFTSVCAGNYMANNTSILNKAKLSFS-----AELSKLPR- 882  
 QY 214 ADCGSLHLIPALNALRNPNN-----SNSSALSILKYSASKKYNISWELGN----- 261  
 DB 883 -----LEKLELNENNLTLQPEINKLRLIYLSVARNKLEISIPDISRLSLKS 931  
 QY 262 ---EPNRYRMEGRAVNSQGLDYQLKSL-----QPIRYSRASLYGPNIGRRKN 312  
 DB 932 LDLSNNRLRMNM---MLEDELTSLVNVSNTLTFGSPAPKFPASPS---PKLAK--- 981  
 QY 313 VTALIDGFMAKVASGVDAVTWQHOCYIDGRVYKVPFLKTRLLD-----TLDQIRKIQV 367  
 DB 982 --SLV--FLSYADNNLTDSIF-----LVNTQNLKTLNLSINNVEISD--LKIQLN 1028  
 QY 368 VNTYTPG-----KQWLEGVYTTSA-----GNNNIS-- 394  
 DB 1029 TELYSQNNFTSLPGEAVOHLRSILKVLMLNKNKLSLPAELSQLSRSLVDVGSNQLKYN 1088  
 QY 395 -DSYAGELVNT-----LGMLANQIDIVYIRHSFPHGVNHLVDQNPFLPDVWLSILY 448  
 DB 1089 ISNHYVDNMNRNNDKLTLPNSGNRRF--IKSLADPEGKNDLSD-----LGIL- 1135  
 QY 449 KRLIPKYLAVHAGLQKRPGRVIRDKLRIYAH-----CTNHN 489  
 DB 1136 KQL---RVLGMDVTLKSKYPPDSVSIRLTTASMINGRVYGVADTLGQSDVCSRDVT 1192  
 QY 490 HNYVGSITLFLINLHRSRKIKYLAQTRDKLVQY-----LLQPYGE-----GLK 536  
 DB 1193 FERFGRDEECILCYDKKNENASGHSKISIRIDYDKLIRLLEKYGESDGIKRALR 1252  
 QY 537 SKSVOLNGO---FLVWVDG 553  
 DB 1253 YSFQLNKEINGMLVSVEDG 1272

## RESULT 10

C88264  
 protein kin-15 [imported] - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 17-May-2002  
 C/Accession: C88264  
 R:anonymous, The C. elegans Sequencing Consortium.  
 Science 282, 2012-2018, 1998  
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology  
 A:Reference number: A75000; MUID:199069613; PMID:9851916  
 A>Note: see websites genome.wustl.edu/gsc/C.elegans/ and ww.sanger.ac.uk/Projects/C.eleg  
 A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
 A:Accession: C88264  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-484 <STO>  
 A:Cross-references: GB:chr\_11; PIDN:CAB01648.1; PID:g3878745; GSPDB:GN00020  
 C:Genetics:  
 A:Gene: kin-15

A:Map position: 2  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

Query Match 3.4%; Score 106; DB 2; Length 484;  
Best Local Similarity 20.7%; Pred. No. 3;  
Matches 89; Conservative 47; Mismatches 150; Indels 144; Gaps 21;

```

QY 62 LLDVSTKN---PV--RTVNEFSLQLDPSI-IHGMWL-----DFLSKRLVTLARGLSPA 111
DB 102 LLDVTSVNEETPIVKPPINERIENLEFDPFEIDQAKLISEDKLGSFGFGEVCYGL--L 159
QY 112 FLRFGKRTDFLOFQNLRNPAKSGGPGPDYLYKNVEDIVRSVDVALDKQKCKIAQHPD 171
DB 160 SMTSTNTEIDTL--QKLSAVAKOSNDPTQENQKMEID-----ETKMCALGRBN 208
QY 172 VMLVLOREKAAQVHLVLEKQFSNTYSNLTIRASLDKYNPADCGSLHIFALNALRRN 231
DB 209 ILAIIQAVTAN-----SGSARNLLIV-----EFVECG----- 235
QY 232 PNNMNSSSALSLIKYSASKKYNISWELGNEPNRYRTMGRAVNGSOLGKDYIQLKSLIQ 291
DB 236 -----DLKFLSEKKSIFKDELVEKNGY-----LL 261
QY 292 PIRIYSPASLYGPNIGRPRNVI-ALLDGMKVA GSTVDAVTWQHCYIDG----- 340
DB 262 PKSIRKTYMENEN-----EDDVIIESLDSL-----CTSDILSPSYQIAEGMEYLASIPCV 312
QY 341 -----RVYKWDPLKTRLDLTSDQIRKIQKVNTYTPGKKIMLEGVTT 386
DB 313 HRDLALRNVLNKKTRIRADFGIARKYQV--DGYRITYGVTGTPMPAR--WMAPEVME 368
QY 387 AGGTNNLSL--SYAAGFLMNTLGLMLANOGI---DVVIRHSFPDHGYNHLVDONFPLPD 441
DB 369 -GKCTEKSDVMYSVGLSYEMFSLGELPYSNVNSDV-----FEHVQGNQOLPMPO 417
QY 442 YMLSILYREL 451
DB 418 YCHPKMYDRM 427

```

## RESULT 11

144330  
protein-tyrosine kinase kin-15 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 09-Mar-1996 #sequence\_revision 09-Mar-1996 #text\_change 17-Nov-2000  
C/Accession: 144330; T23784  
R/Morgan, W.R.; Greenwald, I.  
Mol. Cell. Biol. 13, 7133-7143, 1993  
A:Title: Two novel transmembrane protein tyrosine kinases expressed during Caenorhabditis  
A:Reference number: A54618; MUID:94019384; PMID:8413302  
A/Accession: 144330  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-488 <RES>  
A/Cross-references: GB:103524; NID:g156474; PIDN:AAA28151.1; PID:g156475  
R/Wilkinson, J.  
submitted to the EMBL Data Library, August 1996  
A:Reference number: Z19799  
A/Accession: T23784  
A/Status: preliminary; translated from GB/EMBL/DBD  
A/Molecule type: DNA  
A/Residues: 5-488 <WLI>  
A/Cross-references: EMBL:Z79412; PIDN:CAB01648.1; GSPDB:GN00020; CESP:M176.6  
A/Experimental source: clone M176  
C/Genetics:  
A/Gene: kin-15; CESP:M176.6  
A:Map position: 2  
A/Introns: 29/2; 89/2; 184/1; 226/2; 314/3; 351/2; 427/2; 462/3  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo  
C/Keywords: ATP; phosphotransferase; transmembrane protein  
F:144-459/Domain: protein kinase homology <KIN>  
F:150-158/Region: protein kinase ATP-binding motif

Query Match 3.4%; Score 106; DB 2; Length 488;  
Best Local Similarity 20.7%; Pred. No. 3.1;  
Matches 89; Conservative 47; Mismatches 150; Indels 144; Gaps 21;

```

QY 62 LLDVSTKN---PV--RTVNEFSLQLDPSI-IHGMWL-----DFLSKRLVTLARGLSPA 111
DB 106 LLDVTSVNEETPIVKPPINERIENLEFDPFEIDQAKLISEDKLGSFGFGEVCYGL--L 163
QY 112 FLRFGKRTDFLOFQNLRNPAKSGGPGPDYLYKNVEDIVRSVDVALDKQKCKIAQHPD 171
DB 164 SMTSTNTEIDTL--QKLSAVAKOSNDPTQENQKMEID-----ETKMCALGRBN 212
QY 172 VMLVLOREKAAQVHLVLEKQFSNTYSNLTIRASLDKYNPADCGSLHIFALNALRRN 231
DB 213 ILAIIQAVTAN-----SGSARNLLIV-----EFVECG----- 239
QY 232 PNNMNSSSALSLIKYSASKKYNISWELGNEPNRYRTMGRAVNGSOLGKDYIQLKSLIQ 291
DB 240 -----DLKFLSEKKSIFKDELVEKNGY-----LL 265
QY 292 PIRIYSPASLYGPNIGRPRNVI-ALLDGMKVA GSTVDAVTWQHCYIDG----- 340
DB 266 PKSIRKTYMENEN-----EDDVIIESLDSL-----CTSDILSPSYQIAEGMEYLASIPCV 316
QY 341 -----RVYKWDPLKTRLDLTSDQIRKIQKVNTYTPGKKIMLEGVTT 386
DB 317 HRDLALRNVLNKKTRIRADFGIARKYQV--DGYRITYGVTGTPMPAR--WMAPEVME 372
QY 387 AGGTNNLSL--SYAAGFLMNTLGLMLANOGI---DVVIRHSFPDHGYNHLVDONFPLPD 441
DB 373 -GKCTEKSDVMYSVGLSYEMFSLGELPYSNVNSDV-----FEHVQGNQOLPMPO 421
QY 442 YMLSILYREL 451
DB 422 YCHPKMYDRM 431

```

## RESULT 12

566959  
hypothetical protein YOR076C - yeast (Saccharomyces cerevisiae)  
N/Alternate names: hypothetical protein O2934  
C/Species: Saccharomyces cerevisiae  
C>Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 19-Apr-2002  
C/Accession: S66959  
R/Bohn, C.; Bolotin-Fukuhara, M.; Daignan-Fornier, B.; Dang, D.V.; Valens, M.  
submitted to the Protein Sequence Database, July 1996  
A:Reference number: S66929  
A/Accession: S66959  
A/Molecule type: DNA  
A/Residues: 1-747 <BOH>  
A/Cross-references: EMBL:Z74984; NID:g1420232; PID:e251988; PID:g1420233; MIPS:YOR076C  
A/Experimental source: strain S288C  
C/Genetics:  
A/Gene: SGD:SK17  
A/Cross-references: SGD:S0005602  
A:Map position: 15R  
C:Superfamily: translation elongation factor Tu homology  
C/Keywords: GTP binding; nucleotide binding; P-loop  
F:83-85/Region: GTP-binding SAK/L motif  
F:268-430/Domain: translation elongation factor Tu homology <ETU>  
F:274-281/Region: nucleotide-binding motif A (P-loop)  
F:427-430/Region: GTP-binding NKCD motif

Query Match 3.4%; Score 105; DB 2; Length 747;  
Best Local Similarity 21.9%; Pred. No. 7.1;  
Matches 110; Conservative 66; Mismatches 189; Indels 138; Gaps 26;

```

QY 63 LLDVSTKNP-----VRTVNEFSLQLDPSI-IHGMWLDPLSKRLVTLARGLSPA 115
DB 309 LDPSSNSKAVLVDNTKIREGFSMFKVIQVEN---DLLPSSSTLID--TGSIKY 363
QY 116 GGRRT--DFLOQC-----NLRNPAKSGGPGPDYLYKNVEDIVRSVDVALDKQK 163

```

```

Db      364 FKEKTLNLTEDPEVYVVIDONTDSWEKSLDGNNOY-----ELIKYISYANKNSA 417
Qy      164 CTIAOHPDVMVLQREKA-----ACQELVYLKEQFNTYNSNLTARSLDK 209
      418 CR--KH-----LIILNKADLISMDGRHLEMIOSELNVY-LKENPQMT-----DA 459
Db      210 LYNPADCSGLHLIIPALNLRPNNSWSSSALSILKISA--SKKNI SWELG-----260
Qy      460 BEQFIPCSGL-----LGSNLTENITKTS---KYKSEFDSINYVEMWEGEPFESQ 507
Db      261 -----NEENNRRTMGRAVNSQGLKQYIOLKSLQPIRISRSALSPYNGRPRKNI 314
Qy      508 LVLIVHNNKKTETLEEPVGT-----LLOSSVQPIAEINYSL-----KV 550
Db      315 ALLDGFMYAGSTVDAVWQ--HCY-IDGRVVKVMDIKRLLDLS-----DQIRKIQ 365
Qy      551 LINSGYIQ--SGQTEIHTQDEPHYGIYSRMKNSKQILETWKNNISVGLNPDILLEVLY 609
Db      366 KVVNTYTPGKIM--LEGVTTISAGCTNLSDSYAGFLMTLTIGMLANQIDVIRHSF 423
Qy      610 KIHNTEDFTKQPHIRKGDIIHSRKNTLSPMLP-----NITKULALR-----LIKSI 659
Db      424 FDHYNHLVDONENPLPDYMLSLYKRLIGPKVLAHVAG--LQKRPGRVIRDKRI 480
Qy      660 QTHALSDPVDLQSE-----LLVYHNLTHNAVYKVIQTNDISINPQSLIVEVEIE 712
Db      481 YAHCTHNNHNVYRSGITFIIN 503
Qy      713 PDFALNVDSKITYNNIVLTSID 735

```

## RESULT 13

probable alpha-arabinofuranosidase [imported] - Clostridium acetobutylicum  
 C:Species: Clostridium acetobutylicum  
 C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 30-Sep-2001  
 C:Accession: C97322  
 R:Rolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
 J. Bacteriol. 183, 4823-4838, 2001  
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum  
 A:Reference number: A96900; WUID:21359325; PMID:21359325  
 A:Accession: C97322  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-835 <R>  
 A:Cross-references: GB:AE001437; PIDN:AAK81366.1; PID:G15026526; GSPDS:GN00168  
 A:Experimental source: Clostridium acetobutylicum ATCC824  
 C:Genetics:  
 A:Gene: CAC3436  
 C:Superfamily: Streptomyces alpha-L-arabinofuranosidase SC135.05C

```

Query Match      3.4%; Score 105; DB 2; Length 835;
Best Local Similarity 20.5%; Pred. No. 8.5;
Matches 99; Conservative 61; Mismatches 172; Indels 148; Gaps 24;

Qy      23 APGALYLLALLHH-----SLSSQAGDRPLPVDRAAGLKEKTLILLDVTYKNPVTVE 76
Db      66 ADGGLYSELLONOSFEFKDLSMTVDK-----TGSTST--AEVTSKPLSNNT 114
Qy      77 NPLSLQDPSIIHDGMLDPLSKRLVTLA-RGLSPAVLRFGKRTDPLQFNLRNPAKR 135
Db      115 HYELNCP-----DNNSSLKLVNSGYKGITV--NNAKIDFYFR--ARVNGK--157
Qy      136 GGGPDDYLLKNYEDDIYRSDVALDKQK--CTIAQH-----PDVWLVLQREKAAQWHL 186
Db      158 -GNKVTITQLEDENGNALISEDKITGKINGQWKTEGHLRATKSTSNMAVASITGEAKMN 216
Qy      187 VL-----LKEQPSNTYSNLILTBASLDKLYNFADCSGLH 220
Db      217 DMSVSLFPQDTWKRRKYLKRDIVLRLKDLKPLRFPFGGCIIVGNSKEELIYMKDITIG--274
Qy      221 LIFALNALRBNPN-NSW--NSSSALSLKYSASKKNI SWELQNEP-----NNYRTWEGRA 273

```

```

Db      275 -----NVEERKENTNMGYNQSYGLGFEY-----FQLEDIATVPVLNCGMTQQAIG 324
Qy      274 VNSQGLKDYIOLKSLQPIRISRSALSPYNGRPRKNIALLDGFMYAGSTVAVT- 332
Db      325 VNG-----VBNVAVP-----GPDIDPIYQNAVLD-----VEYANGDASTY 360
Qy      333 WQCYITGRVYKVMDFPKTLDDT-----LSQIIRKIQVNVVYTPGKKIMLWEGVTT 386
Db      361 WGRKRISGKHKEPN-LKYVALIGNEQMPYHHRFEAFQVNLQCKP-----GITLIS 412
Qy      387 AGTNNLSDSYAAGFLMTLTIGMLANQIDVIRHSFDDHYNHLVDONENPLPDYML 445
Db      413 NAGTSPSGSTFDDNMWIK-----EKAPTVVDEHYMSPDWFLS 452

```

## RESULT 14

MG422 homolog C12.orf839 - Mycoplasma pneumoniae (strain ATCC 29342)  
 S73548  
 C:Species: Mycoplasma pneumoniae  
 A:Variety: ATCC 29342  
 C:Date: 27-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 07-Dec-1999  
 C:Accession: S73548  
 R:Himmelreich, R.; Halbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Hermann, R.  
 Nucleic Acids Res. 24, 4420-4449, 1996  
 A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae  
 A:Reference number: S73327; WUID:97105885; PMID:8948633  
 A:Accession: S73548  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-839 <H>  
 A:Cross-references: EMBL:AE000022; GB:U00089; NID:G1673882; PIDN:AA895870.1; PID:G1673882  
 A:Note: The nucleotide sequence was submitted to the EMBL Data Library, November 1996  
 C:Genetics:  
 C:Superfamily: hypothetical protein MG422

```

Query Match      3.4%; Score 104.5; DB 2; Length 839;
Best Local Similarity 20.2%; Pred. No. 9.3;
Matches 109; Conservative 85; Mismatches 181; Indels 165; Gaps 27;

Qy      70 PVRATNENF--LSIQD-----PSII-HDGMLDPLSKRLVTL--ARGSPAFVRFQGR 119
Db      374 PQOSLNMKTIQAISELEKTLTPYVLTDAVIGFLKRLAFLQRYAGLARC--EQNKL 431
Qy      120 TDFLOFQNLNRNPAKRGPGPDYIKN---YEDDIVRSDVALDKQKCKIAQHPDVMV 175
Db      432 VNELKQNVLLNQNGYKGEVSQSYVNVQKAVFKFIQTVKAAALNYTK-LKHTLDPEFLMN 490
Qy      176 LQREKAAQWHLVYLKEQPSNTYSNLILTBASLDKLYNF--ADCSGLH-----220
Db      491 IVQERCF-----EQLLTYYEKLDWTKELNQLYVCRALNMLKQTOHFTTKYQFI 542
Qy      221 ----LIFALN-----ALRRPNNSWSSSALSILKYSASKKNI SWELQNEPNRY 267
Db      543 THGVVDFFVNGQGRNQTQASLKAHLNPDWNSPKKLLVNVKYNKYFEANL--HHPAYL 599
Qy      268 TMEGRAVNSQGLK-----DYQLKS-----LQPIIYRSALSYGNIGRPR 310
Db      600 LIPNR--NATLTETNTTTINQMLKTQKQRAHYHLLODIDILIO--WYKKEIKQKQ 655
Qy      311 KVVALLDGFMYAGSTVDAVWQHCYIDGIVKMDPELKTRLDLSLQIRKIQKYVNT 370
Db      656 QQIKRLKLY-----GTLNKL-----LNTQISKYNNV--683
Qy      371 YTPGKKIWLWEGVTTISAGCTNLSDSYAGFLMTNT-----LGMLAN 412
Db      684 ----RKTFF--VSEECDLNRLOASNDLHFNLLNANVNVISFCIKCRONPKKLNTNAN 736
Qy      413 QGIDVIRHSFDDHYNHLVDONENPL--PDVWLSLYKRLIGPKVLAHVAGLQKRP 469
Db      737 --LKKVLDNTRFKNGIPSMWIFSDLNKINTKQRFKLYLLFKLLHPLQVLLV-----784

```

QY 470 PGRVIRDKLRIYAHCTNNHNNVYRGSTITLFINLRSRKIK-LAGTLRDKLVHGYLLQ 528  
DB 785 -----DSFVNPF-----NKTITVNTFRG-----LLIANQNOGIAVLFPNDPHNNLVKDFFTQ 629

## RESULT 15

B34087  
hypothetical protein (L1H 3' region) - human  
C/Species: Homo sapiens (man)  
C/Date: 30-Mar-1990 #sequence\_revision 30-Mar-1990 #text\_change 30-Sep-1993  
C/Accession: B34087  
R/Scott, A.F.; Schmeckpeper, B.J.; Abdelrazik, M.; Comey, C.T.; O'Hara, B.; Rossiter, J.  
Genomics 1, 113-125, 1987  
A/Title: Origin of the human L1 elements: proposed progenitor genes deduced from a conserved  
A/Reference number: A34087; MUID:88085185; PMID:3692483  
A/Accession: B34087  
A/Status: preliminary; not compared with conceptual translation  
A/Molecule type: DNA  
A/Residues: 1-1280 <SCO>  
C/Superfamily: pol polyprotein

Query Match 3.4%; Score 104.5; DB 2; Length 1280;

Best Local Similarity 19.4%; Pred. No. 18;  
Matches 127; Conservative 94; Mismatches 216; Indels 217; Gaps 34;

QY 50 DRAAGLEKTLILDDVSTKNPVRT-----VNEPFLIQDPSIHDGWLDFLSKRL 101  
DB 136 DLQRDLDSHTIINGDEFNT--PLSTLRSTRQKNKDIQELN---SALHQ-----DL 182  
QY 102 VTLARGISPAFLFGCGRTDPLQFNLRNPAKSGRGPVYLNKVEDDIVRSDVALDQ 161  
DB 183 IDIYRTLHP-----KSTERYTF-----SAPHHTYSKIDHILSKALLSKC 222  
QY 162 KGCKI-----AQHPDVLVLQREKAAQM-----LVL-----LKEQFSN----- 195  
DB 223 KRTETIITNCISDSHSAICLRLIKLTONHSTWKLNNLINDYVHNEMKAEIKNPFETN 282  
QY 196 -----TYSNLIITARSJDKYNPADCGMLIFALNALRNPNNSMNSSALSLLKYSAS 250  
DB 283 ENKDTTYQNLMDTFKAV-----CRGKFT--ALNNAHK-----KQERS 317  
QY 251 KKYNISWELNGNEPNRYRTMGRVNSQLGKDYIOLKSL-----LOPIRIYSRSLYGP- 304  
DB 318 KIDPLTSQI--KELEKQOTSKASRROETIKIRALKEIETQKLOKIN--ESRSWPFEXI 375  
QY 305 -NIGRPKNVIA-----LDDGPKVAGS--TVDAVWQHCHYIDGRVVKWDFIKTRLDDI 356  
DB 376 NKIDRPLARLIKKGKGNQIDSIKNDKGDITDPTETIQ-----TTIREYKHLIYANK 427  
QY 357 LSDQIRKIQKVVNTYT-----PGKIMLEGVVTTSAGTNNLSDSYAAG 400  
DB 428 L-ENLEMDKFLDLYTLPRLNQEEVESLNRPITGSEI--EATINSLEPTKKSPPDGPFTAE 484  
QY 401 F-----LMLNTLGMLANOGIDVIVIRHSPFD-----HGYNHLVDQNFNPLPDY 442  
DB 485 FYGRYKEELVPFLKLPQSIKEGI---LPSFYEASIIILPKGRDYYTKENFRPIS-- 539  
QY 443 WLSLLYRRLIGPKVLA-----VHVAGLQKRPGRV-----IRDKLRIYAHCTNNHNNH 490  
DB 540 -LNMIDAKILN-KILANQIQCHIKLILHNOVGRIIPGQGMENIRKSINVIGHINRTXK 597  
QY 491 NYVRGSTITLINLRSRKIKILAGTLRDKLVHGYLLQYGOBGL-----K 536  
DB 598 NHM-----IISID-----AEKAFDKIQOPFMLKTNKLGIDGTYLKIRAIYDKP 642  
QY 537 SKSYQNGQPLVWVDDGTLPELKPRLPAGTTLVIPPVTMGFYVVKVNNALACR 590  
DB 643 TANIILNGQ-----KLEAFPLKGTGRGCPILSPILFNIVLEVLARATR 685

Search completed: May 6, 2004, 13:46:13  
Job time : 22 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2004, 13:41:21 ; Search time 18 Seconds  
without alignments  
1712.530 Million cell updates/sec

Title: US-10-088-676-2

Perfect score: 3088

Sequence: 1 MRLVLCAPPEAMPSSNSRPPA.....PTMGFYVKNVNALACRRY 592

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	123	4.0	2026	1 CYAA_YEAST	P08678 saccharomyc
2	115	3.7	923	1 YEL4_YEAST	P40017 saccharomyc
3	114.5	3.7	837	1 XYNZ_CLODM	P10478 clostridium
4	106.5	3.4	1839	1 CYAA_SACKL	P23466 saccharomyc
5	106	3.4	488	1 K115_CAEEL	P34891 caenorhabdi
6	104.5	3.4	839	1 Y422_MYCPN	P75175 mycoplasma
7	104	3.4	6359	1 BACC_BACLI	O68008 b bacitraci
8	101	3.3	535	1 COX1_HANWI	P48868 hansenula w
9	101	3.3	779	1 LON_BRECH	P67772 brevitacill
10	101	3.3	2971	1 VXC5_CHRE	O32065 chlamydomon
11	100.5	3.3	630	1 GIDA_TREPA	O83084 treponema p
12	100.5	3.3	870	1 SUL2_HUVAN	O81w50 homo sapien
13	100.5	3.3	878	1 MSH4_YEAST	P40965 saccharomyc
14	100.5	3.3	1259	1 LINI_HUVAN	P08547 homo sapien
15	100.5	3.3	3343	1 YOG7_CAEEL	P34616 caenorhabdi
16	99	3.2	546	1 SYR_THEAC	O9147 thermoplasma
17	99	3.2	871	1 SCL0_YEAST	O06245 saccharomyc
18	98.5	3.2	836	1 GKI1_MOUSE	O60934 mus musculu
19	98.5	3.2	1356	1 ROM2_YEAST	P51862 saccharomyc
20	98	3.2	629	1 T3MH_HABIN	P71366 haemophilus
21	98	3.2	866	1 RECE_ECOLI	P15032 escherichia
22	97	3.1	543	1 YDE3_SCHPO	O10437 schizosacch
23	97	3.1	897	1 DPOL_HELPJ	O92159 helicobacte
24	96.5	3.1	544	1 RGSE_RAT	O08773 rattus norv
25	96	3.1	4036	1 RPL1_DUGBY	O66431 dugbe virus
26	95.5	3.1	709	1 STL1_CAEEL	O41336 caenorhabdi
27	95.5	3.1	2136	1 YCF2_MARPO	P09975 marchantia
28	95	3.1	2492	1 TAL4_DICDI	P44633 dictyosteli
29	94.5	3.1	695	1 NIC3_DROME	O9VC27 dirosophila
30	94.5	3.1	832	1 ANR3_HUVAN	P57240 buchiera ap
31	94.5	3.1	870	1 SUL1_MOUSE	P57240 buchiera ap
32	94.5	3.1	1235	1 KPB2_HUVAN	P20365 kuytremomyc
33	94.5	3.1	1235	1 KPB2_MOUSE	P49007 alteromonas

34	94	3.0	625	1 GIDA_IACIA	O9CEJ4 lactococcus
35	94	3.0	641	1 LIP_STAYH	P04635 staphylococ
36	94	3.0	839	1 N19E_YEAST	P34077 saccharomyc
37	94	3.0	1517	1 RPOC_CAME	O9P130 campylobact
38	93.5	3.0	870	1 SUL1_RAT	O9V160 rattus norv
39	93.5	3.0	1106	1 C1C2_RABIT	P13806 oryctolagus
40	93	3.0	1075	1 PST2_SCHPO	O13919 schizosacch
41	93	3.0	1300	1 POL2_MOUSE	P11369 mus musculu
42	92	3.0	430	1 SURA_BUCAI	P57240 buchiera ap
43	92	3.0	461	1 GP13_YEAST	P23263 saccharomyc
44	92	3.0	534	1 COX1_KLITA	P20365 kuytremomyc
45	92	3.0	773	1 HEXB_ALTSO	P49007 alteromonas

## ALIGNMENTS

RESULT 1  
CYAA\_YEAST STANDARD; PRT; 2026 AA.  
ID CYAA\_YEAST  
AC P08678;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Adenylate cyclase (EC 4.6.1.1) (ATP pyrophosphate-lyase) (adenyllyl cyclase).  
GN CYR1 OR CDC35 OR HSR1 OR SRA4 OR YJL005W OR J1401.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_Taxid=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86079531; PubMed=293438;  
RA Kataoka T., Broek D., Wiggler M.;  
RT "DNA sequence and characterization of the S. cerevisiae gene encoding adenylate cyclase."  
RL Cell 43:493-505(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA To Van D., Perea J., Jacq C.;  
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / FY1679;  
RA de Haan W., Smits P.H.M., Griwell L.A.;  
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE OF 1042-2026 FROM N.A.  
RX MEDLINE=88165073; PubMed=3327602;  
RA Masson P., Lenzen G., Jacquemin J.M., Danchin A.;  
RL "Yeast adenylate cyclase catalytic domain is carboxy terminal."  
RN Curr. Genet. 10:343-352(1986).  
RN [5]  
RP MUTAGENESIS OF THR-1651.  
RX MEDLINE=9112042; PubMed=1991451;  
RA Feger G., de Vendittis E., Vitelli A., Masturzo P., Zahn R., Verotti A.C., Kavounis C., Pal G.P., Fasano O.;  
RL "Identification of regulatory residues of the yeast adenyllyl cyclase."  
RN EMBL J. 10:349-358(1991).  
RL -1- FUNCTION: Plays essential roles in regulation of cellular metabolism by catalyzing the synthesis of a second messenger, cAMP.  
CC -1- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.  
CC -1- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).  
CC -1- ENZYME REGULATION: The presence of GTP-bound Ras2 protein is required in order to elicit a magnesium-dependent adenyllyl cyclase activity.  
CC -1- SIMILARITY: Belongs to the adenyllyl cyclase class-3 family.  
CC -1- SIMILARITY: Contains 21 leucine-rich (LRR) repeats.  
CC -1- SIMILARITY: Contains 1 P2C-like domain.  
CC -1- SIMILARITY: Contains 1 Ras-associating domain.



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
CC EMBL; ML2057; AAA34549.1; -;  
CC EMBL; Z49280; CAA89295.1; -;  
CC EMBL; X87611; CAA60917.1; -;  
CC EMBL; X03448; CAA27175.1; -;  
CC PIR; S56776; OMBY.  
CC Germodline; 141621; -;  
CC SGD; S0003542; CYR1.  
CC GO; GO:0005886; C:Plasma membrane; IDA.  
CC GO; GO:0007126; P:metosis; IMP.  
CC GO; GO:0000074; P:regulation of cell cycle; IMP.  
CC InterPro; IPR001054; G:cyclase.  
CC InterPro; IPR001611; LRR.  
CC InterPro; IPR003591; LRR\_cyp.  
CC InterPro; IPR001932; PP2C-like.  
CC InterPro; IPR000159; RA domain.  
CC Pfam; PF00211; Guanylate\_cyc; 1.  
CC Pfam; PF00560; LRR; 15.  
CC Pfam; PF00481; PP2C; 1.  
CC SMART; SM00044; C1CC; 1.  
CC SMART; SM00332; PP2C; 1.  
CC SMART; SM00314; RA; 1.  
CC PROSITE; PS50125; GUANYLATE\_CYCLASES\_2; 1.  
CC PROSITE; PS50200; RA; 1.  
CC Lysase; Repeat; Leucine-rich repeat; cAMP biosynthesis; Metal-binding;  
CC Magnesium.  
CC KX  
CC DOMAIN 676 755 PAS-ASSOCIATING.  
CC REPEAT 815 838 LRR 1.  
CC REPEAT 842 862 LRR 2.  
CC REPEAT 863 885 LRR 3.  
CC REPEAT 886 908 LRR 4.  
CC REPEAT 910 931 LRR 5.  
CC REPEAT 932 955 LRR 6.  
CC REPEAT 957 976 LRR 7.  
CC REPEAT 977 999 LRR 8.  
CC REPEAT 1001 1016 LRR 9.  
CC REPEAT 1017 1040 LRR 10.  
CC REPEAT 1042 1062 LRR 11.  
CC REPEAT 1063 1086 LRR 12.  
CC REPEAT 1088 1109 LRR 13.  
CC REPEAT 1110 1132 LRR 14.  
CC REPEAT 1134 1156 LRR 15.  
CC REPEAT 1188 1209 LRR 16.  
CC REPEAT 1210 1232 LRR 17.  
CC REPEAT 1233 1256 LRR 18.  
CC REPEAT 1258 1280 LRR 19.  
CC REPEAT 1285 1308 LRR 20.  
CC REPEAT 1319 1344 LRR 21.  
CC REPEAT 1369 1625 PRPC-LIKE.  
CC DOMAIN 1626 2026 CATALYTIC.  
CC METAL 1673 1673 MAGNESIUM (BY SIMILARITY).  
CC METAL 1716 1716 MAGNESIUM (BY SIMILARITY).  
CC MUTAGEN 1651 1651 T->I: WEAK PAS-INDEPENDENT ACTIVITY.  
CC MUTAGEN 1651 1651 T->D: ATTENUATION OF THE RESPONSE TO  
CC RAS PROTEINS.  
CC FT 262 262 S->L (IN REF. 1).  
CC FT 548 548 S->L (IN REF. 1).  
CC FT 592 592 D->H (IN REF. 1).  
CC FT 709 709 R->I (IN REF. 1).  
CC FT 962 962 L->P (IN REF. 1).  
CC FT 1388 1388 E->S (IN REF. 1).  
CC FT 1427 1427 E->D (IN REF. 3).  
CC FT 1461 1461 A->S (IN REF. 1).  
CC FT 1566 1566 A->S (IN REF. 1).  
CC FT 1735 1735 V->G (IN REF. 1).  
CC FT 1735 1735

FT CONFLICT 1956 1956 I -> V (IN REF. 3).  
FT CONFLICT 1996 1996 C -> P (IN REF. 1).  
FT CONFLICT 2009 2026 NVDELQKAKKADIST -> MLTNFYKLTQRTQYOLE  
FT FT  
SQ SEQUENCE 2026 AA; 227832 MW; BEAB7419AB6989D0 CRC64;  
Query Match 4.0%; Score 123; DB 1; Length 2026;  
Best Local Similarity 22.6%; Pred. No. 1.1; Indels 162; Gaps 33;  
Matches 141; Conservative 83; Mismatches 237;  
QY 37 LSSQAGRRPLPVDAAGLKE-KTLLIDVSTKNPV---RTVNEFLSLQDPSTTHDQW 92  
DB 891 LNLQNELESIP-----AGFVELKNLQLLDDSSNKFMYPEVINVCNLLQIDISYKIQS 946  
QY 93 LDFLSKRLVTLAR-GLSPAFLREGGRTPD--LQFQNLK-----NPAKSGGPGP 140  
DB 947 LP-QSTKYLVLAKONLSHNKLNFTGLSEKTDRTLNKYNRLSITKTAASLQNLFLT 1005  
QY 141 DYLLKNEYEDIVRSVDALDKQGGCKIAQHPDVMVLQREKAAQW-HLVLLKEQPSNTYSN 139  
DB 1006 DNRISNFPD-----TLPKLRALEIQENFTISIFSDFPKXNTSLTLNKAQLSSIPGE 1058  
QY 200 LILPARSIDKYNPADSGGLHITALLALRNPNMNSNSSLKLYKSAXKYNISMEI 259  
DB 1059 LILKLSFLEKLE-----LNQNNLTFLQEI-SKLTKLVLFSVANKKLEYLPEL 1106  
QY 260 GNEPNNYRT--MGRVAVNSQLGQDIYOLKSLQPIRIYRASD-----YGPNIQR 308  
DB 1107 -SQLKSRLTDLDHNNIRDPYDGMENELTSLNLSNAPNSLSLNSFYHNSYGSKLKSK 1165  
QY 309 PRKNVIALDQPMKAYGSTDAVWQ--HCYIDGRVYKVMDFLKTLLDLSQIRIQK 366  
DB 1166 ----SLW--FPLAQNQFDAMPLENFVN---LKVNLSTYNSFDV---SHKLES 1211  
QY 367 VVNTYTPGKXIMLGSVTVTSAGTNNLSDSYAAGFLN--INTLGLMLNQGIDV-----V 418  
DB 1212 ITELYSNKL-----TTLSGDT-----VLKMSLSKTLMLNSQNLSPALSLNL 1256  
QY 419 IRHSFDFGVNHLVDQNFPLPDY-WLS--LKYRLIGP----- 454  
DB 1257 SQLSVFDVGANQLKKNISNHYDMNMRNKKELKYNFSGNRFEIKSPISHDIDADISDL 1316  
QY 455 -----KYLAVHVAAGLQKPRPGRVIRDKLRIYACTNNHNN-----NYV----- 493  
DB 1317 TWLPQKLKVLGMDVTINTKTPDENVAFRLKTSIINGKRYGVAADTLGGRDVYSRDVT 1376  
QY 494 ---RGSITLPLINRSRKX-----IKLAGTLADK-LVHOYLLQPYGQF----- 533  
DB 1377 FEFRRGNDECLCLCHDSKNONADYGNISIRIVDIYDKILIRQ--LERYGDETDNIKT 1434  
QY 534 GLKSKSVOLNQG---ELVMVVDG 553  
DB 1435 ALRFSFLQNKELINGMLNSVDNG 1457  
RESULT 2  
YIELD YEAST STANDARD; PRF; 923 AA.  
AC P40017;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hypochemical 103.3 kDa protein in PRO3-GCD11 intergenic region.  
GN YER024W.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxId=4932;  
RN 11  
SEQUENCE FROM N.A.  
RC STRAIN=6286C / AB972;  
RX MEDLINE=97313264; PubMed=9169868;  
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Yelton M.A., Allen E.,

RA Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,  
 RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,  
 RA Huntke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,  
 RA Lin D., Mosedale D., Nakahara K., Nameth A., Norgren R., Oeffner P.,  
 RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,  
 RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.,  
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.",  
 RL Nature 387:78-81(1997).  
 CC -1- SIMILARITY: Belongs to the carnitine/choleline acetyltransferase  
 CC family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: U18778; AAB64557.1; -.  
 DR PIR: S50482; S50482.  
 DR GerMOnline: 139104; -.  
 DR SCD: S0000826; YER024W.  
 DR GO: GO:0004092; F:carnitine O-acetyltransferase activity; IMP.  
 DR GO: GO:0006066; P:alcohol metabolism; IGI.  
 DR GO: GO:0009437; P:carnitine metabolism; IMP.  
 DR InterPro: IPR000542; Carn\_acyl\_trans.  
 DR Pfam: PF00755; Carn\_acyltransfer\_1; PALSE\_NEG.  
 DR PROSITE: PS00439; ACYLTRANSF\_C\_1; PALSE\_NEG.  
 DR PROSITE: PS00440; ACYLTRANSF\_C\_2; 1.  
 KM Hypothetical protein; Transferase; Acyltransferase.  
 SO SEQUENCE 923 AA; 103333 MW; B59AB881491D68A7 CRC64;  
 Query March 3.7%; Score 115; DB 1; Length 923;  
 Best Local Similarity 18.7%; Pred. No. 1.4; Indels 206; Gaps 27;  
 Matches 115; Conservative 79; Mismatches 216;  
 33 LHSLSQAGDRRLPYDRAAGLEKTLILLDSTKNPVTNENFLS-LQDPS--II 88  
 98 LHDINNQTSRE-----IQDDVLPNPFVLADDLLENITQDRSAVIV 142  
 89 HDGWLDSLSEKRLVTLARG--LSPAFIRFGKRTDFIQFQNL--RNPAKSGG----- 137  
 143 H-----SARFISALKQDLLPDPINATNGKPLSMAPFLNFGTTRSPVFGSEVENFD 135  
 138 -----PGPDY-----YLNKYEDIVRSQVALLDKQKQCKIAQHPDVMVLVQ 177  
 196 LNKPYTASDLEDPEYSSDEDNDPEPQKDPDRKRKHEEDIPTNGNITIKRHPDSKILL 255  
 178 REKAQWELVLKQFNSNTYSLILTRASLDKLYN--PDCSGHLIFALNAL----- 228  
 256 ISRGQVYTLLEV-----DSTWKIYTAALTTTFNHIITKSSGEEKSTALSGTSHSFRN 310  
 229 -----RNPNNWNSSSALSLKYASKKYINWELGNEPNNRYTHGAVNGSQL 279  
 311 WKYARKLQKYPVLELRIDSLAFVLDSEQETTN--DGDYADISQMFNRTITRDX 368  
 280 GKDYIOLKSLLOPTIRIVSRASLYGPNIGRPKNVIALLDGFMKYGSTVDA---VTWQHC 336  
 369 KCTSANCK-----RVFYGTSIINSKGQVSCVSRWYDKQLQV--TADAKATVWDSF 420  
 337 YIDGRVYVWDELKTRLLDLSQDIRKQKQVNVYTPGKKIM-----LEGVV 383  
 421 TCGDSV--VLAFTS---EITYESVRLARLVNAGDQFSLMPVNTQMDPETHKLMATYI 474  
 384 TTSAGTNNISDSYAAGFLMNTLGLMANQGDIVVIRHSPFDHGYNLVQNNENPLPDYV 443  
 475 SADGGSPEDIPK-----LVNKK-----IWSFS----- 498  
 444 LSLLYKLLIGPKLVAVHVGQKRPGRVRLDKLRIYAHCTNHNHNVYRGSTTLFIIN 503  
 499 -----NINLTHVHLETK-----IADLISKY-----DIYASAIPL----- 528

QY 504 LHSRKXIKIAGTRDLYHQVY-----LQPYGQEGK-SKSVQINQPLV 548  
 DB 529 ---GRSAQRIGVQPDWVQALQAIHAYLXGRWVGLFVSTRGFNSNSPTINISQA 585  
 QY 549 MVD-----DGT 554  
 DB 586 LLELCQLFISSIDGT 601  
 RESULT 3  
 XYNZ CLOTM STANDARD; PRT; 837 AA.  
 ID XYNZ CLOTM  
 AC P10478;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Endo-1,4-beta-xylanase Z precursor (EC 3.2.1.6) (Xylanase Z)  
 GN (1,4-beta-D-xylan xylanohydrolase Z).  
 OS Clostridium thermocellum.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 OC NCBI\_TaxId=1515;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NCIB 10682;  
 RX MEDLINE=8908072; PubMed=3139632;  
 RA Grepinet O., Chebrou M.-C., Beguin P.;  
 RT "Nucleotide sequence and deletion analysis of the xylanase gene  
 RT (xynZ) of Clostridium thermocellum.",  
 RL J. Bacteriol. 170:4582-4586(1988).  
 RN (2)  
 RP X-RAY CRYSTALLOGRAPHY (1.4 ANGSTROMS) OF 515-837.  
 RC STRAIN=NCIB 10682;  
 RX MEDLINE=9539342; PubMed=7664125;  
 RA Dominguez R., Souchon H., Spinelli S., Dauteur Z., Wilson K.S.,  
 RA Chauvaud S., Beguin P., Alzari P.M.;  
 RT "A common protein fold and similar active site in two distinct  
 RT families of beta-glycanases".  
 RL Nat. Struct. Biol. 2:569-576(1995).  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-xylosidic  
 CC linkages in xylans.  
 CC -1- DOMAIN: A 24 RESIDUES DOMAIN IS REPEATED TWICE IN THIS ENZYME AS  
 CC WELL AS IN OTHER C.THERMOCCELLUM CELLULOSE ENZYMES. THIS DOMAIN  
 CC MAY FUNCTION AS THE BINDING LIGAND FOR THE SL COMPONENT.  
 CC -1- SIMILARITY: Belongs to cellulase family F (family 10 of glycosyl  
 CC hydrolases).  
 CC -1- SIMILARITY: Contains 1 xynZ-type cellulose-binding (CBD) domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: M22624; AAA23286.1; -.  
 DR PIR: A31842; A31842.  
 DR PDB: 1XYZ; 29-JAN-96.  
 DR PDB: 1JUF; 31-OCT-01.  
 DR PDB: 1J72; 27-MAR-02.  
 DR InterPro: IPR006584; CBD IV.  
 DR InterPro: IPR002105; CBD 6.  
 DR InterPro: IPR002105; Dockerin\_1.  
 DR InterPro: IPR002048; EF-hand.  
 DR InterPro: IPR000801; Sesterase\_put.  
 DR InterPro: IPR008979; Gal\_bind\_like.  
 DR InterPro: IPR001000; Glyco\_hydro\_10.  
 DR InterPro: IPR000379; Ser\_esters.  
 DR Pfam: PF00422; Dockerin\_1; 2.  
 DR Pfam: PF00404; Dockerin\_1.  
 DR Pfam: PF00756; Esterase\_1.

DR Pfam; PF00331; Glyco\_hydro\_10; 1.  
 DR PRINTS; PR00134; GLYHYDRLASE10.  
 DR SMART; SM00606; CBD IV; 1.  
 DR SMART; SM00633; Glyco\_10; 1.  
 DR PROSITE; PS00018; EF\_HAND; UNKNOWN 2.  
 DR PROSITE; PS00448; CLOS\_CELULOSOME\_RPT; 2.  
 DR PROSITE; PS00591; GLYCOSYL\_HYDROL\_F10; 1.  
 KM Xylan degradation; Hydrolase; Glycosidase; Repeat; Signal;  
 3D-structure.  
 FT SIGNAL 1 28  
 FT CRAIN 29 837  
 FT ACT\_SITE 645 645  
 FT ACT\_SITE 754 754  
 FT DOMAIN 328 416  
 FT DOMAIN 430 487  
 FT REPEAT 430 453  
 FT REPEAT 464 487  
 FT DISULFID 783 789  
 FT HELIX 518 524  
 FT TURN 525 526  
 FT STRAND 528 533  
 FT TURN 535 535  
 FT HELIX 536 539  
 FT TURN 540 540  
 FT HELIX 543 552  
 FT STRAND 553 558  
 FT TURN 561 562  
 FT HELIX 564 567  
 FT STRAND 570 570  
 FT TURN 571 572  
 FT STRAND 573 573  
 FT HELIX 577 588  
 FT TURN 589 590  
 FT STRAND 592 599  
 FT HELIX 606 609  
 FT TURN 610 610  
 FT HELIX 615 632  
 FT TURN 633 636  
 FT STRAND 639 644  
 FT STRAND 647 647  
 FT TURN 649 650  
 FT STRAND 654 654  
 FT HELIX 658 663  
 FT TURN 665 666  
 FT HELIX 667 678  
 FT TURN 680 681  
 FT STRAND 683 688  
 FT HELIX 696 710  
 FT TURN 711 712  
 FT STRAND 717 720  
 FT STRAND 723 725  
 FT HELIX 730 745  
 FT TURN 746 747  
 FT STRAND 749 760  
 FT TURN 761 762  
 FT HELIX 765 785  
 FT TURN 787 788  
 FT STRAND 789 794  
 FT STRAND 798 799  
 FT TURN 800 807  
 FT HELIX 804 809  
 FT TURN 810 810  
 FT STRAND 811 812  
 FT STRAND 817 817  
 FT TURN 819 820  
 FT STRAND 823 823  
 FT HELIX 825 834  
 FT SEQUENCE 837 AA; 92262 MW; DD4C29F04D12B6CD CRC64;  
 Query Match 3.7%; Score 114.5; DB 1; Length 837;  
 Best Local Similarity 21.0%; Pred. NO. 1.4; Indels 107; Gaps 20;  
 Matches 77; Conservative 54; Mismatches 128;

QY 128 LRNPAGKRG---GPGDYIYLNKTEDDIVASDVALDKOKGCKINQHPDWLVLOREKAAQM 184  
 DB 518 LRDYAEARGIKIGTCVNPFPYNSDPYNS-----LIQREFSMVVCENE----- 561  
 QY 185 HLVLTKQFSNTVSNLLTLTASLDKYNFADCSGL-----HLIFALNALRANP-----NSW 236  
 DB 562 ---MKEDALOPRQNVDFSKG-DQLAFERRGMQREGHTLTHN---QNPMLNNGW 613  
 QY 237 NSSSALSILKXAS-----KRYNISWELANE-----PNRYRTMAGRAVNGSLGXDYIQ 285  
 DB 614 NRDSILAWMKHITVTHWYKKGIVEMDVANECWDSGNGLRSSIMRWV---IGDYLD 669  
 QY 286 LKSLLOPRIRYSAS-----LVG-----PNIQPRKXVIALLOGFMKVASSTVAVWQH 335  
 DB 670 YAFR-----YAFRADPDALLFTVDVIEDLG-PKSNVAVNMKSKVREVPIDGVGFQC 722  
 QY 336 CYIDGRVYKVDPLKRLDPLDSPDIRKIQV-----VNTYTPGKKIMLGQVTTSA 387  
 DB 723 HFING-----NSPEYIASIDQNIKRYAIEIGVIVSFTFDIRIPQS---ENPATAFQ 770  
 QY 388 GGTNNLSQSYAGFLMTLNLGLANQSIDVIVRSFDPH-----GYNH--LVQNF 436  
 DB 771 VQANNYKELMKI-----CLANPNCNFFVMWGFDTKYTWLPGTPPGNPLTIDSNY 821  
 QY 437 NPLPDY 442  
 DB 822 NPKPAY 827

RESULT 4  
 CYAA\_SACKL ID CYAA\_SACKL STANDARD; PRT; 1839 AA.  
 AC P23466;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Adenylyl cyclase (EC 4.6.1.1) (ATP pyrophosphate-lyase) (Adenylyl cyclase).  
 GN CYR1.  
 OS Saccharomyces kluyveri (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxId=4934;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91323718; PubMed=1864503;  
 RA Young D., O'Neill K., Broek D., Wigler M.;  
 RL "The adenylyl cyclase-encoding gene from Saccharomyces kluyveri.";  
 RL Gene 102:125-132(1991).  
 CC -!- FUNCTION: Plays essential roles in regulation of cellular  
 CC metabolism by catalyzing the synthesis of a second messenger,  
 CC cAMP.  
 CC -!- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.  
 CC -!- COFACTOR: Binds 1 magnesium ion per subunit (by similarity).  
 CC -!- SIMILARITY: Belongs to the adenylyl cyclase class-3 family.  
 CC -!- SIMILARITY: Contains 21 leucine-rich (LRR) repeats.  
 CC -!- SIMILARITY: Contains 1 P2C-like domain.  
 CC -!- SIMILARITY: Contains 1 Ras-associating domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

DR EMBL; X56042; CAA39513.1; -.  
 DR PIR; J01145; OYBYK.  
 DR InterPro; IPR001054; G\_cyclase.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR003591; LRR\_Typ.  
 DR InterPro; IPR001992; P2C-like.

DR InterPro: IPR00159; RA\_domain.  
 DR Pfam: PF00211; guanylate\_cyc; 1.  
 DR Pfam: PF00560; LRR 15.  
 DR Pfam: PF00481; PP2C; 1.  
 DR SMART: SM00044; CYCC; 1.  
 DR SMART: SM00369; LRR\_TYF; 2.  
 DR SMART: SM00332; PP2C; 1.  
 DR SMART: SM00314; RA; 1.  
 DR PROSITE: PS50125; GUANYLATE\_CYCLASES\_2; 1.  
 DR PROSITE: PS50200; RA; 1.  
 DR Lyase: Repeat; Leucine-rich repeat; CAMP biosynthesis; Metal-binding; Magnesium.  
 KM Lyase: Repeat; Leucine-rich repeat; CAMP biosynthesis; Metal-binding; Magnesium.  
 FT DOMAIN 494 574 PAS-ASSOCIATING.  
 FT REPEAT 632 655 LRR 1.  
 FT REPEAT 659 679 LRR 2.  
 FT REPEAT 680 702 LRR 3.  
 FT REPEAT 703 724 LRR 4.  
 FT REPEAT 726 748 LRR 5.  
 FT REPEAT 749 771 LRR 6.  
 FT REPEAT 773 793 LRR 7.  
 FT REPEAT 794 818 LRR 8.  
 FT REPEAT 820 833 LRR 9.  
 FT REPEAT 834 857 LRR 10.  
 FT REPEAT 859 879 LRR 11.  
 FT REPEAT 880 903 LRR 12.  
 FT REPEAT 905 926 LRR 13.  
 FT REPEAT 927 949 LRR 14.  
 FT REPEAT 951 974 LRR 15.  
 FT REPEAT 1004 1025 LRR 16.  
 FT REPEAT 1026 1048 LRR 17.  
 FT REPEAT 1050 1072 LRR 18.  
 FT REPEAT 1074 1096 LRR 19.  
 FT REPEAT 1101 1124 LRR 20.  
 FT REPEAT 1135 1160 LRR 21.  
 FT REPEAT 1185 1440 PP2C-LIKE.  
 FT DOMAIN 1441 1839 CATALYTIC.  
 FT DOMAIN 1488 1488 MAGNESIUM (BY SIMILARITY).  
 FT METAL 1531 1531 MAGNESIUM (BY SIMILARITY).  
 SQ SEQUENCE 1839 AA; 206895 MW; 86A69BCB1F2733CB C64;  
 Query March 3.4%; Score 106.5; DB 1; Length 1839;  
 Best Local Similarity 21.3%; Pred. No. 16;  
 Matches 132; Conservative 75; Mismatches 220; Indels 193; Gaps 33;  
 56 KEKTLILLDVSTK---NPRTVNE--NFLSLQDPSIIHDGMDLFLSKKLVTLARGLSP 110  
 724 KTKKVLQLLDITISSNKFVYEVYINSCNMLQIDLSYKIKHS--LP-VSINQVLVLAKE--- 776  
 111 AFLRFGGKRT---DFLQFQNR-----NPAKSRGGGP-----DYLLKNYEDIVR 153  
 777 -KNLENNRLTSVGLSQPKNRLTLNRCNRYTSIECHAFNLQNLPLTDNRISTDDDLTR 835  
 154 SDVALDKQKCKIAQHPDVMVLVLERAAQWHLVLEQPSNTYSLILTAASIDKLYNF 213  
 836 -----LRTLELQNPITSWGCGANWAMNTSLINKAKLSRPS-----AEILSKLPR- 882  
 214 ADGCGHLIPLNALRRPNP-----SNSSSALSLIKTSASAKKNISWELG----- 261  
 883 -----LEKLELNENNLTLQPEINLTLIYLSVARNKLASIPEDISLRSLSKS 931  
 262 -----EPNNYRTWGRAVNGSQKDYIOLKSL-----QPIRTYSRASLYGPNIGRPKN 312  
 932 LDHSHNNLRLMNM--NLEDLELSLNVSSNLLGFHSGPAKFPASP-----PKLAK----- 981  
 313 VIALIDGPMKAGSTVDAVWQHCYIDGRVYKWMDFLKTLLD-----TISDQIRKIQV 367  
 982 --SL--FLSYADNNLTDSIWP-----LVNTEQNKTLNLSYNNFEVLSID--LKLQNL 1028  
 368 VNTYTPG-----KTIWLEGVVTSAG-----GTNNLS-- 394  
 1029 TELYLSGNFTSLGSAVQHRLSKVLTMLNKNKLSLIPAEISQSLRSLYDVGSNQKYN 1088  
 395 -DSYAAAGFLWINT-----LGMLANOGIDIVIRHSFPDHGYNHLYVDQNFNPLDYWLSLLY 448

DB 1089 ISNHYVDMNRKNTDCLKLTFSGNKRFE--IKSLDPBGKNDLSD-----LELL- 1135  
 QY 449 KLLIGPKYLAHVAVAGLQKRPGRGVIRDKLRIYAH-----CTNHN 489  
 DB 1136 KQL---RVLGMDVTLKTSKVPDESVSIRLRTTASMINGMKRYGAVDTLGQSDSVCSRDT 1192  
 QY 490 HNYVAGSTLTPIINLHRRKKIKLAGIRDLVQY-----LLQPYGQ-----GLK 536  
 DB 1193 FERFGRDEICILCYDCKNNASSGKHSKIRIDYDKIIRULEKYGESDGIKRALR 1252  
 QY 537 SKSVQINGC--PLVMVDDG 553  
 DB 1253 YEFQLANEINGMKLVSVSDG 1272  
 RESULT 5  
 ID K115 CAEEL STANDARD; PRT; 488 AA.  
 AC P34891;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DE 15-MAR-2004 (Rel. 43, Last annotation update)  
 GN KIN-15 OR M176.6.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=94019384; PubMed=84113302;  
 RA Morgan W.R., Greenwald I.;  
 RT "Two novel transmembrane protein tyrosine kinases expressed during  
 RT Caenorhabditis elegans hypodermal development.";  
 RL Mol. Cell. Biol. 13:7133-7143(1993).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Wilkerson J.;  
 RL Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.  
 CC - FUNCTION: May be specifically involved in cell-cell interactions  
 CC regulating cell fusions that generate the hypodermis during  
 CC postembryonic development. It has a role in the development of the  
 CC HYP7 hypodermal synyrium.  
 CC - CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC - TISSUE SPECIFICITY: Hypodermal cells.  
 CC - DEVELOPMENTAL STAGE: Expressed during hypodermal development.  
 CC - SIMILARITY: Belongs to the Tyr family of protein kinases.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL collection -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L03524; AAA28151.1; -  
 DR EMBL: 278412; CAB01648.1; ALT\_INIT.  
 DR PIR: I44330; I44330.  
 DR HSSP: P11362; 1FGK.  
 DR WormBep: M176.6; CE12470.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR InterPro: IPR008266; Tyr\_kinase\_AS.  
 DR Pfam: PF00069; tyknae; 1.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR ProDom: PD000001; Prot\_kinase; 2.  
 DR SMART: SM00219; Tyrc; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.

DR PROSITE, PS00109; PROTEIN KINASE TYR; 1.  
 DR PROSITE, PS0011; PROTEIN\_KINASE\_DOM; 1.  
 KW Transferase; Tyrosine-protein kinase; ATP-binding; Receptor;  
 KW Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 26  
 FT CHAIN 27 488 RECEPTOR-LIKE TYROSINE-PROTEIN KINASE  
 FT DOMAIN 27 488  
 FT TRANSMEM 51 70 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 51 488 POTENTIAL.  
 FT DOMAIN 144 458 CYTOPLASMIC (POTENTIAL).  
 FT NP BIND 150 158 PROTEIN KINASE.  
 FT BINDING 183 183 ATP (BY SIMILARITY).  
 FT ACT SITE 319 319 ATP (BY SIMILARITY).  
 FT CAREGND 25 25 BY SIMILARITY.  
 SQ SEQUENCE 488 AA; 56544 MW; F891B9BEA128977 CRC64; N-LINKED (GLCNAC... (POTENTIAL).

Query Match 3.4%; Score 106; DB 1; Length 488;  
 Best Local Similarity 20.7%; Pred. No. 2.7;  
 Matches 89; Conservative 47; Mismatches 150; Indels 144; Gaps 21;

QY 62 LLDVSTK--P--RTVNNFLSLQDPSI-IHDGML---DFLSKRLVTLARGLSPA 111  
 DB 106 LLDVSTVNEETPIVKKPPIKERIENLEDFEIDQAKLEISBDKLSGFGEGVCGY--L 163  
 QY 112 FLRFGGKRTDPLQFOLRNPAKSGRGPDPYLYKNEYDIDVRSDVALDKQKCKIAQHPD 171  
 DB 164 SMTSTMTEDTL--QLTSVAVKOSNPPTQENCKMTED-----ETKMCALIGRNPN 212  
 QY 172 VMLVLRKRAQAHVLVLEKQPSNTYSNLTILTRASLDKLYNPADCSGLHIFALNALRN 231  
 DB 213 ILAIIGAVTN-----SGSARNLLIV-----EFVECG----- 239  
 QY 232 PNNSSNSSALSLTKSASKYNIISWELGEPNNRTMGRAVNSQLGKDYIQLKSLDQ 291  
 DB 240 -----DLKFLKEKKSIFKDELYENKGY-----LL 265  
 QY 292 PIPIYSRSLYGNIGRPKNVI-ALIDGPMKAGSTVDAVTWQHCYIDG----- 340  
 DB 266 PKSIIRKTYFMEN-----EDDVIEESIDSL-----CTSDLSFSYQIAGMEYLASIPCV 316  
 QY 341 -----RVYKQVDFLKTBLDPLSDQIRKIOKVNTYTPRGKIMLEGVVTS 386  
 DB 317 HRDLARNVTLNOKTIRIDPGLARKYQ--DGYRITKGVGTMPMPAR--WMAPEVARE 372  
 QY 387 AGGTNNLSD--SYAAGFLWNTLQMLANQI---DVVIRHSFPGHGYNHLVDONFNPLD 441  
 DB 373 -GKTEKSDVMSYGVSLYEMFSLGELPYSNVNSDV-----FEHVQGNQLPMPQ 421  
 QY 442 YMLSLYKRL 451  
 DB 422 YCHPKMIDRM 431

RESULT 6  
 Y422\_MYCPN STANDARD; PRT; 839 AA.  
 ID AC P51575;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 GN Hypothetical protein MG422 homolog (C12\_orf839).  
 DE MN620 OR MP222.  
 OS Mycoplasma pneumoniae.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2104;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 29342 / M129;  
 RX MEDLINE=97105885; PubMed=8948633;  
 RA Himmelfreisch R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,  
 Hermann R.;  
 "Complete sequence analysis of the genome of the bacterium Mycoplasma

RT pneumoniae";  
 RL Nucleic Acids Res. 24:4420-4449 (1996).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: AE000022; AAB95870.1; --  
 DR PIR: S73548; S73548.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 839 AA; 99912 MW; 3E353E18BCDFECC CRC64;

Query Match 3.4%; Score 104.5; DB 1; Length 839;  
 Best Local Similarity 20.2%; Pred. No. 7.5;  
 Matches 109; Conservative 85; Mismatches 181; Indels 165; Gaps 27;

QY 70 PRTVNEP--LSQLD-----PSTI-HDGLDPLSSRLVTL--ARGSPAFLREGGR 119  
 DB 374 PGLSLNMYQATISFELEKTRPLVTHDAVIGFLYSRLFLQRYAKLNRC--EQHNKL 431  
 QY 120 TDFLOFOLRNPAKSGRGPDPYLYKNEYDIDVRSDVALDKQKCKIAQHPDMLV 175  
 DB 432 VNELKQNVTLNQYKGEVSQSVVQKAVFGFIQTVKALNTYK-LKHTLDPFLNM 490  
 QY 176 LQREKAQAHVLVLEKQPSNTYSNLTILTRASLDKLYNF--ADCSGLH----- 220  
 DB 491 IVQERCP-----EQLITTYEKLDTKYLQNLVFCRALMTNLHKQTQHFYTKQFI 542  
 QY 221 --LIFALN-----ALRRPNNSSNSALSLTKSASKYNIISWELGEPNNRTMG 267  
 DB 543 THGVVDVFNQGRNQTPASLKNLNRDNMSPKMLLVNKTWVYFANL--HHQAYL 599  
 QY 268 TMAGRAVNSQLGK-----DYQLKS-----LLQPIRYSRSLYGNIGRP 310  
 DB 600 LLEPNR--NATTLTEANTTTINQNLTKQKWRAPHYLLQDRIILQ--WLYKEIKQKQ 655  
 QY 311 KNVIALDGEFMKAGSTVDAVTWQHCYIDGRVYKQVDFLKTBLDPLSDQIRKIOKVNT 370  
 DB 656 QQIKALKNY-----GTNLKL-----LNTQISKVNNVV-- 683  
 QY 371 YTPGKATWLBGVVTTSGGTTNNTSDSYAAGFLWNT-----LQMLAN 412  
 DB 684 ---RKTFF--VDSECDNRLQASNKLFNLLNMAVNIISFCLKKCRQNPCKLRTAN 736  
 QY 413 QGIDVIRHSFPGHGYNHLVDONFNPL--PDYWSLYKRLIGPKYLAHVAVGLQKRP 469  
 DB 737 --LQMLDNTKQKIPKMIPLSDLNKNTYQKRFYLLFLGLHPQVLV----- 784  
 QY 470 PGRVIRDLRIYAHCTNHNHNHYVRSITLFIINLHRSRKIK-LAGTLRDKLVHGYLLQ 528  
 DB 785 -----DSFVNF-----NKHVTNFTRG-----LILAHQNOGIIAYLFENDPHNLVKKQFFTQ 829

RESULT 7  
 BACC\_BACLI STANDARD; PRT; 6359 AA.  
 ID AC O68008;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 GN Bacitracin synthetase 3 (Bac3) [includes: ATP-dependent isoleucine  
 DE adenylation (IleA) (Isoleucine activase); ATP-dependent D-phenylalanine  
 DE adenylation (D-PheA) (D-phenylalanine activase); ATP-dependent histidine  
 DE adenylation (HisA) (Histidine activase); ATP-dependent D-aspartate  
 DE adenylation (D-AspA) (D-aspartate activase); ATP-dependent asparagine  
 DE adenylation (AsnA) (Asparagine activase); Aspartate racemase  
 DE (EC 5.1.1.13); Phenylalanine racemase [ATP hydrolyzing]  
 DE (EC 5.1.1.11)].  
 GN BACC.

CC Bacillus licheniformis.  
 CC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 CC NCBI\_TaxID=1402;  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC STRAIN=ATCC 10716;  
 CC MEDLINE=96089193; PubMed=9427658;  
 CC Konz D., Klems A., Schoegendorfer K., Marahiel M.A.,  
 CC "The bacitracin biosynthesis operon of Bacillus licheniformis ATCC  
 CC 10716: molecular characterization of three multi-modular peptide  
 CC synthetases.";  
 CC Chem. Biol. 4:927-937(1997).  
 CC -1- FUNCTION: INDUCES PEPTIDE SYNTHESIS, ACTIVATES AND INCORPORATES  
 CC FIVE AMINO ACIDS, FORMS A THIAZOLINE RING BETWEEN THE FIRST TWO  
 CC AMINO ACIDS AND INCORPORATES A D-GLUTAMINE IN THE FOURTH POSITION.  
 CC -1- CATALYTIC ACTIVITY: L-aspartate = D-aspartate.  
 CC -1- CATALYTIC ACTIVITY: ATP + L-phenylalanine = AMP + diphosphate + D-  
 CC phenylalanine.  
 CC -1- COFACTOR: Contains 5 covalently bound phosphopantetheines  
 CC (potential).  
 CC -1- PATHWAY: Cyclic peptide antibiotic bacitracin biosynthesis.  
 CC -1- SUBUNIT: LARGE MULTISUBUNIT COMPLEX OF BAI, BAZ AND BAC.  
 CC -1- DOMAIN: CONSISTS OF FIVE MODULES WITH TWO EPIMERIZATION DOMAINS IN  
 CC THE SECOND AND FOURTH MODULES, AND A PUTATIVE C-TERMINAL  
 CC THIOESTERASE DOMAIN. EACH MODULE INCORPORATES ONE AMINO ACID INTO  
 CC THE PEPTIDE PRODUCT AND CAN BE FURTHER SUBDIVIDED INTO DOMAINS  
 CC RESPONSIBLE FOR SUBSTRATE ADENYLATION, THIOATION, CONDENSATION  
 CC (NOT FOR THE INITIATION MODULE), AND EPIMERIZATION (OPTIONAL), AND  
 CC N METHYLATION (OPTIONAL).  
 CC -1- MISCELLANEOUS: BACITRACIN IS A MIXTURE OF AT LEAST TEN CYCLIC  
 CC DODECAPEPTIDES, THAT DIFFER BY ONE OR TWO AMINO ACIDS. THE MOST  
 CC ABUNDANT IS BACITRACIN A, A BRANCHED CYCLIC DODECAPEPTIDE. IT  
 CC CONTAINS AN N-TERMINAL LINEAR PENTAPEPTIDE MOIETY (ILE-CYS-LEU-D-  
 CC GLU-ILE) WITH AN ISOLEUCINE-CYSTEINE THIAZOLINE CONDENSATION  
 CC PRODUCT AND A C-TERMINAL HEPTAPEPTIDE RING (LYS-D-ORN-ILE-D-PHE-  
 CC HIS-D-ASP-ASN), IN WHICH THE FREE ALPHA-CARBOXY GROUP OF THE C-  
 CC TERMINAL ASN IS BOUND TO THE EPSILON-AMINO GROUP OF LYSINE. IT  
 CC CONTAINS FOUR AMINO ACIDS IN THE D-CONFIGURATION (GLU-4, ORN-7,  
 CC PHE-9, AND ASP-11).  
 CC -1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme  
 CC family.  
 CC -1- SIMILARITY: Contains 5 acyl carrier domains.  
 CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 CC EMBL: AF007865; AAC06348.1; -  
 CC PIR: T31679; T31679.  
 CC HSRP: P14687; 1AMU.  
 CC InterPro: IPR000873; AMP-bind.  
 CC InterPro: IPR001242; Condensacn.  
 CC InterPro: IPR006163; PP-bind.  
 CC InterPro: IPR006162; Pantne\_S.  
 CC InterPro: IPR001031; Thioesterase.  
 CC Pfam: PF00501; AMP-binding; 5.  
 CC Pfam: PF00668; Condensation; 7.  
 CC Pfam: PF00550; PP-binding; 5.  
 CC Pfam: PF00975; Thioesterase; 1.  
 CC PRINTS: PR00154; AMPBINDING.  
 CC PROSITE: PS0012; PHOSPHOPANTHETHEINE, 4.  
 CC PROSITE: PS00455; AMP\_BINDING, 5.  
 CC PROSITE: PS0075; ACP DOMAIN, 5.  
 CC Liaser, Isomerase, Hydrolyase; Antibiotic biosynthesis;  
 CC Phosphopantetheine; Multifunctional enzyme; Repeat.  
 CC REPEAT 461 1034 DOMAIN 1 (ISOLEUCINE-ACTIVATING).  
 CC REPEAT 1517 2064 DOMAIN 2 (D-PHENYLALANINE-ACTIVATING).  
 CC REPEAT 2999 3570 DOMAIN 3 (HISTIDINE-ACTIVATING).  
 CC REPEAT 4047 4612 DOMAIN 4 (D-ASPARTIC ACID-ACTIVATING).  
 CC

FT REPEAT 5549 6129 DOMAIN 5 (ASPARAGINE-ACTIVATING).  
 FT 966 1033 ACYL CARRIER (ACP) 1.  
 FT DOMAIN 1998 2064 ACYL CARRIER (ACP) 2.  
 FT 3502 3569 ACYL CARRIER (ACP) 3.  
 FT DOMAIN 4544 4610 ACYL CARRIER (ACP) 4.  
 FT 6052 6119 ACYL CARRIER (ACP) 5.  
 FT BINDING 996 996 PHOSPHOPANTHETHEINE (BY SIMILARITY).  
 FT BINDING 2028 2028 PHOSPHOPANTHETHEINE (BY SIMILARITY).  
 FT BINDING 3532 3532 PHOSPHOPANTHETHEINE (BY SIMILARITY).  
 FT BINDING 4574 4574 PHOSPHOPANTHETHEINE (BY SIMILARITY).  
 FT BINDING 6082 6082 PHOSPHOPANTHETHEINE (BY SIMILARITY).  
 SQ SEQUENCE 6359 AA; 722923 MW; 82A273C546253074 CRC64;  
 Query Match 3.4%; Score 104; DB 1; Length 6359;  
 Best Local Similarity 19.6%; Pred. No. 1.4e+02;  
 Matches: 88; Conservative 71; Mismatches 163; Indels 128; Gaps 17;  
 QY 20 ACTAPGATVYALLHLSSQAGDRPLPVDRAGLKEKTLILDVSTKNPVTVNENFL 79  
 DB 3680 APTKPALEKAPLELAAYVAKASGDH-----VLMDMH----- 3712  
 QY 80 SLQDPSITIDGKMLDFLSKRLVTLARG--LSPALRFGGKRTDPLQONTL--RNPAS 134  
 DB 3713 -----HIISDVSAIFRELAELYEGLTPEPLTYQ---KDFSEWQKLFYQDEVR 3762  
 QY 135 RGGSPGPDYLYKQVEDIVASVDALDKQCKIAQHPDWLV-----LQREKAA 182  
 DB 3763 Q-----EDYVNLNFGSEVPLNIPADEKRPQKSGISGDIVQFIDETSAMINKAKENGA 3818  
 QY 183 QMHLVLKEQFNTSYNLTITARSLSDKYNFADCSGLH-----LIFALNALRNPVN 233  
 DB 3819 TWYMLLA-----GYTTLAKTGGEDIVGSPIGRHHSDIKHYIGFINTLAVRNHRK 3873  
 QY 234 NSNNSSSALSLKYASAKKYNISWELGSPNNRYRMHGAVVGSLQKDYIOLKSLQPI 293  
 DB 3874 GDMFPDYLYKEVKEITLKY-----ENQDY-----PDELYEKLQVYK 3910  
 QY 294 RYVSRAIVPNIGPRKRVNVALDGFMAKAGSTVDAVTWQHCYIDGRVYKMDLKTRL 353  
 DB 3911 RDMSRRLPFDVTL-----VLQNPFG-----DEADIDGLTFQPLQRE-----VNISKEDL 3954  
 QY 354 LDTLSQIRKIQKVVNTVYFGKIVLEGVYTSAGG--TNISDSVAAAGTMTNGLMAN 412  
 DB 3955 TLTAATNEGICVFYST--KLFRSTTERWAGHLINILKEAANDPMPPLSDVNMUS- 4010  
 QY 413 QGIDVYIRHSFPHGYNHLVDQNFPLDPY 442  
 DB 4011 -----DEEMNALDDQNGKQADY 4028  
 RESULT 8  
 COXI\_HANWT STANDARD; PRT; 535 AA.  
 ID COXI\_HANWT  
 AC P48668;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 GN Cytochrome c oxidase polypeptide I (Ec 1.9.3.1).  
 OS COXI.  
 OS Hansenula winei (Yeast).  
 OC Mitochondrion.  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Pichia.  
 CC NCBI\_TaxID=4907;  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC STRAIN=21;  
 CC Sekito T., Okamoto K., Kitano H., Yoshida K.;  
 CC "Yeast Hansenula winei mitochondria genome's complete DNA sequence  
 CC demonstrated unique characteristics";  
 CC Nucleic Acids Symp. Ser. 31:233-234 (1994).  
 CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory  
 CC chain that catalyzes the reduction of oxygen to water. Subunits

CC 1-3 form the functional core of the enzyme complex. CO I is the  
 CC catalytic subunit of the enzyme. Electrons originating in  
 CC cytochrome c are transferred via the copper A center of subunit 2  
 CC and heme A of subunit 1 to the binuclear center formed by heme A3  
 CC and copper B.  
 CC -1- CATALYTIC ACTIVITY: 4 ferrocyanochrome c + O(2) = 4 ferrocyanochrome  
 CC c + 2 H(2)O.  
 CC -1- PATHWAY: Respiratory chain; terminal step.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
 CC inner membrane. Contains 12 potential transmembrane domains.  
 CC -1- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: D31785; BAA06563.2; -.  
 CC PIR: S58740; S58740.  
 CC HSSP: P98002; 1AR1.  
 CC InterPro: IPR000883; COX1.  
 CC Pfam: PF00115; COX1; 1.  
 CC PRINTS: PRO0077; COX1; 1.  
 CC PROSITE: PS00077; COX1; 1.  
 CC Oxidoreductase; Heme; Copper; Mitochondrion; Transmembrane;  
 CC Respiratory chain; Inner membrane.  
 CC FT METAL 63 63 IRON (HEME A AXIAL LIGAND) (PROBABLE).  
 CC FT METAL 242 242 COPPER B (PROBABLE).  
 CC FT METAL 246 246 COPPER B (PROBABLE).  
 CC FT METAL 291 291 COPPER B (PROBABLE).  
 CC FT METAL 292 292 COPPER B (PROBABLE).  
 CC FT METAL 377 377 IRON (HEME A3 AXIAL LIGAND) (PROBABLE).  
 CC FT METAL 379 379 IRON (HEME A AXIAL LIGAND) (PROBABLE).  
 CC FT METAL 379 379 IRON (HEME A AXIAL LIGAND) (PROBABLE).  
 CC FT METAL 242 246 1'-histidyl-3'-tyrosine (His-Tyr)  
 CC CROSSLINK  
 CC  
 CC FT SEQUENCE 535 AA; 59025 MW; 97F7C4EADAD50A CRC64;  
 CC  
 CC SQ  
 CC  
 CC Query Match 3.3%; Score 101; DB 1; Length 535;  
 CC Best Local Similarity 22.2%; Pred. No. 7.3;  
 CC Matches 68; Conservative 46; Mismatches 102; Indels 90; Gaps 19;  
 CC  
 CC QY 293 IRIYSASLYGPNIGRRKRVVILLDGPKVAGSTDAVW-OHCYIDGVVVMDFLKT 351  
 CC DB 260 VSTYSKKVFGGE-----ISNVAMASIA-FLGFLVSHMTIVG-----LD-ADT 302  
 CC QY 352 RLDDTSDQIRKIQKVVNTYTPGKVLWEGVVTTSAGTNNLSDS--YAGFLMNTLG- 408  
 CC DB 303 RAVFTSTWVIAVPTGKIFS-----WL-----ATLVGGSIRLAVPMVLALFLFTIGG 353  
 CC QY 409 -----MLANQGDVIRISFPDHGYNHY--DONENPLPRYW-----LSLYKR----- 450  
 CC DB 354 LTVGVALNANSLDAVHDYTVVGGFHYVLSWGAIFSLFAGYVWSPQILGLVYNERLAQT 413  
 CC QY 451 -----LIGPKV--LAVVAGLQKRPGRVIRDKLRIYAHCTHHNNHYRG----- 495  
 CC DB 414 QFWLIFGAVVIFPMHFLGQGMKPR-----RIPDPDAVAGNVTSSIGVYAI 464  
 CC QY 496 SITLFTLNHSRKKIKLAGTLRDKLVHQ--YLLOP-----YQDELKSKSVQ--L 542  
 CC DB 465 SLAFIYIYID-----OLINGLTNKKDKSVVSKADFEVSNITFANNISKASIEFLL 519  
 CC QY 543 NGQPLV 548  
 CC DB 520 NSPPAI 525  
 CC  
 CC RESULT 9  
 CC LON BRECH STANDARD; PRT; 779 AA.  
 CC ID LON BRECH  
 CC AC P36772;

DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE ATP-dependent protease La (EC 3.4.21.53).  
 GN LON.  
 OS Brevibacillus choshinensis.  
 OC Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Brevibacillus.  
 OK NCBI\_TaxID=54911;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HPD31;  
 RA MEDLINE=92202157; Pubmed=1551846;  
 RA Ito K., Ueda S., Yamagata H.;  
 RT "Cloning, characterization, and inactivation of the Bacillus brevis  
 RL lon gene.";  
 CC J. Bacteriol. 174:2281-2287(1992).  
 CC -1- FUNCTION: Degrades short-lived regulatory and abnormal proteins in  
 CC presence of ATP. Hydrolyzes two ATPs for each peptide bond cleaved  
 CC in the protein substrate (By similarity).  
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of large proteins such as globin,  
 CC casein and denaturated serum albumin, in presence of ATP.  
 CC -1- SIMILARITY: Belongs to peptidase family S16.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: D00863; BAA00737.1; -.  
 CC MEROPS: S16.001; -.  
 CC InterPro: IPR003593; AAA ATPase.  
 CC InterPro: IPR003959; AAA ATPase, centr.  
 CC InterPro: IPR001270; Chaprinin, clpA/B.  
 CC InterPro: IPR008269; Pept. S16-C.  
 CC InterPro: IPR004815; Pept. S16-Lon.  
 CC InterPro: IPR003111; Pept. S16-N.  
 CC InterPro: IPR008268; Peptid. S16-AS.  
 CC InterPro: IPR001984; Peptidase\_S16.  
 CC Pfam: PF00004; AAA; 1.  
 CC Pfam: PF02190; LON; 1.  
 CC Pfam: PF03362; Lon C; 1.  
 CC PRINTS: PRO0100; CLP/PROTEASE.  
 CC PRINTS: PRO0830; ENDOLAPTASE.  
 CC SMART: SMO0382; AAA; 1.  
 CC SMART: SMO0464; LON; 1.  
 CC TIGR/Pfam: TIGR00763; Lon; 1.  
 CC PROSITE: PS01046; LON\_SER; 1.  
 CC Hydrolyase; Serine protease; ATP-binding.  
 CC FT NP BIND 355 362 ATP (POTENTIAL).  
 CC FT ACT SITE 678 678 BY SIMILARITY.  
 CC SEQUENCE 779 AA; 87421 MW; 625B0DA9B98941B5 CRC64;  
 CC  
 CC SQ  
 CC  
 CC Query Match 3.3%; Score 101; DB 1; Length 779;  
 CC Best Local Similarity 19.8%; Pred. No. 12;  
 CC Matches 126; Conservative 80; Mismatches 181; Indels 250; Gaps 33;  
 CC  
 CC QY 9 EAMPSNSRPPACAPGAL---YLALLHLHSLSSQAGRRRLPVDRAAGLKEKTLILDDV 65  
 CC DB 278 EAMPSNS-----AESGVIRITTYDTLFLAPWITTEEDN-----LDI 312  
 CC QY 66 STKNPVRVNEFNFLSLQDPSIHDGWLDFLSKRLVYLAAG-----LSP 110  
 CC DB 313 --KHAEEVLDEHDYGLERKERV---LEVLAQKLVMSKGPILCVGPPGVKTSILAR 366  
 CC QY 111 AFLRFGGRIDFLOQNFNAPKSRG-----PG-----PDYLYKVED 149  
 CC DB 367 SVARALGREFVRIISIGVRDELRGRRRYVAGLPGRIIOGKQAGTINVFLLD--EI 424  
 CC QY 150 DIVRSDV-----ALDRQGGCKIAQHPDVMLVLOREKPAQWHLVLKQFSNTYSN 199



```

Db      425 DKLASFRGDPASALLEVDPNQNDKFSDH-----XIEETYLITNMW 466
Qy      200 LITLAFSLKLVNFAQSCGHLIFALNALRRNNPNNSSALSLKYSKKYNI--SW 257
Db      467 FITTANSLLTI-----PRLLDMEVYISIGYTELEKNIIRGY 505
Qy      258 ELGNEBNRYTRMGRAVNSQLGKYIOLK--SLLOPIRIYRASLYG--PNIGRPKNYI 314
Db      506 LEPKQMED--HG-----LQKDKLQNMEDAMKLVLTYREA--GVANINREANVC 552
Qy      315 ----ALLDQPMK---VAGSTVAVWQHCYIDGRVYVMDFLKRLDLTLDQIRKQK 366
Db      553 RKAATIVGGEKKRVVVTATLLEALGKPRYGLAEK-----KQVGSV-- 597
Qy      367 VVNTYTPGKKIMLEGVTTYSAGG--TNNLSDSYAAGFLMNTLGLMLAN-----QGIDVVI 419
Db      598 -----TGLAWTQAGDGLNVEVSIAGKGLTLTGQDGVKESQAALFSYI 644
Qy      420 RHSEFFDHGYNHLYDQNFNPDPYWSLILYKRLIGPKYLAHV--AGLQKRPGRVIRDKL 478
Db      645 RSRASEWG---IDPEFHEKND-----IHIVEGAIIPKDGPSAGITMAT 685
Qy      479 RIRAHCTNHNHNYVRSITLFTIINLRSRKIKLAG--TLRDKLVHQYLLQPYGQGLK 536
Db      686 ALVSALT-----GIPV-----KKEVGMGTETLRGRVL-----PIG--GLK 719
Qy      537 SKSVQLNGQPLVWVDGTLPELKPRLPAGRTLVIRP 573
Db      720 EKCMSAH-----RAGLITITILP 736

```

## RESULT 10

ICX3\_CHLRE STANDARD; PRT; 2971 AA.

```

AC Q32065; Q95635;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical 341.7 kDa protein in pebd-pbdc intergenic region
DE (ORF2971) (ORFB).
OS Chlamydomonas reinhardtii.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_Taxid=3055;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=137c / CC-125;
RA Watson A.T., Puton S.;
RT "Unidentified open reading frame ORF2971 (ORFB) from the chloroplast
RL genome of Chlamydomonas reinhardtii."
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 2635-2971 FROM N.A.
RC STRAIN=137c / CC-125;
RX MEDLINE=2305394; PubMed=12417694;
RA Maul J.E., Lilly J.W., Cul L., dePamphilis C.W., Miller W.,
RA Harris E.H., Steen D.B.;
RT "The Chlamydomonas reinhardtii plastid chromosome: islands of genes in
RL a sea of repeats."
RL Plant Cell 14:2659-2679 (2002).
RN [3]
RX MEDLINE=81013-1021(1989).

```

## COMPLETE PLASTID GENE.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

Query Match 3.3%; Score 101; DB 1; Length 2971;  
Best Local Similarity 20.7%; Pred. No. 81;  
Matches 127; Conservative 87; Mismatches 207; Indels 192; Gaps 34;

```

Qy      62 LLDVSTKQPVRYNE--NPLSLQDLPSTIHDQWDLFSSKRLVT---LARGSPAFARFG 116
Db      306 LLD--SKKATQTPSLYFVFTQLNQKTFSLYMLPVAGLALITPTLLTLTGQSVYQKN 363
Qy      117 ---GRTDFLOFQNLNRPASRG-----GPGDYLYKNY-----EDIVRSQVALDKQ 161
Db      364 SFINKTDMVLSNTEWPSKSGFTPTLFGTSVEIYLPNSYWPKEGSGGINRVSSINAV 423
Qy      162 KGCKIAGHPDVMVLQREK-----AAQMHVYLKEQPSNTYSNLI--LTARSLDLQYFAD 215
Db      424 KKVATYAN--LVLDSEQEVATSFQNDLISICVNNLNYNSKNTALSTKRLFLFSA 479
Qy      216 CSGL-----HLIPALNLRN--PNNSSNSALSGLK 246
Db      480 IKSNAIKHKTQSFSEVNTTLLGNNSPVYKGFKSSINAFSSYLPSTNVHSMPLTSLP 539
Qy      247 Y--SASKKYN-----ISHELGNP-----NNYRTMG--RAVNGS 277
Db      540 YLKALISPLYSKFMIDHSIKFTPTTKLQHKLNKSPKQNTYTKQFTGLRDLALNSF 599
Qy      278 QLGKQYIO---LKSLOPIRIRYRASLYGNIGRPKNYIALDGEPMKV-----AGS 326
Db      600 SFGQVNFRTNHFILNSNRPLAHVNA-----LKLINGEYQYQNNLQINCNK 645
Qy      327 TVDAVWQHCYIDGRVYVMDFLKRLDLTSDQIRKIQKQVNTYTGKKIMLEGVTT 386
Db      646 TLD-----LNRKNKLQVYVHKSHEFNQKSOIVYKQSLYN-----BDLCIR 687
Qy      387 AGGTNNLSDSYAAGFLMNTLGLMLANGIDVIVIRHSFFDHGYNHLYD---QENPPLPDY 442
Db      688 GNGT--KVVDYFSHGDKLSNKGIVLDY---FVSNLLEFNTNTIINKDGKQNTYKX-- 741
Qy      443 WLSTL---YKRLIGPYLAHV--VAGLQKRPGRVIRDKLR1-YAHCTNH---HNH 490
Db      742 -LNLKTTVPKTYLI--KRYTSINSIVANEQ-----TRNNLWGIHFNGHSVVSNA 791
Qy      491 NVVRGSITLFTIINLRSRKIKLAGTLDKLVHQYLLQPYGQGLKSKSVQALQCPVWV 550
Db      792 NLLTG-----RPVKFYIKRKDKRLNSYLI--YVNOULK--KFIDLNNN----- 830
Qy      551 DDGTLPELKPRL 563
Db      831 -----FLKPKPL 837

```

## RESULT 11

GIDA\_TREPA STANDARD; PRT; 630 AA.

```

ID GIDA_TREPA
AC O83084;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 26-FEB-2003 (Rel. 41, Last annotation update)
DE Glucose inhibited division protein A.
GN GIDA OR TP0044.
OS Treponema pallidum.

```

CC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.  
 OX NCBI\_TaxId=160;  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=Nichols;  
 RX MEDLINE=98332770; PubMed=9665876;  
 RA Fraser C.W., Norris S.J., Weinstock G.M., White O., Sutton G.G.,  
 RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,  
 RA Dodesgren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,  
 RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,  
 RA McDonald L., Arriach P., Bowman C., Cotton M.D., Fujii C., Garland S.,  
 RA Hatch B., Horst K., Roberts K., Sandusky M., Weiman J., Smith H.O.,  
 RA Venter J.C.;  
 RT "Complete genome sequence of Treponema pallidum, the syphilis  
 RT Spirochete";  
 RL Science 281:375-388(1998).  
 CC -1- FUNCTION: Not known.  
 CC -1- SIMILARITY: Belongs to the gda family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL, AE001189; AAC65038.1; -.  
 DR F1R; C71374; C71374.  
 DR TIGR; TP0044; -.  
 DR HAMAP; MF 00129; -1.  
 DR InterPro: IPR001327; FAD pyr\_redox.  
 DR InterPro: IPR002218; GIDA.  
 DR InterPro: IPR004415; GIDA\_sub.  
 DR Pfam; PF01134; GIDA\_1.  
 DR PRINTS; PR00368; FADPFR.  
 DR ProDom; PD003738; GIDA\_1.  
 DR TIGRFAMs; TIGR00136; gida\_1.  
 DR PROSITE; PS01280; GIDA\_1; 1.  
 DR PROSITE; PS01281; GIDA\_2; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 630 AA; 70175 MW; 1B52C6F02CEC275 CRC64;  
 Query Match 3.3%; Score 100.5; DB 1; Length 630;  
 Best Local Similarity 21.3%; Pred. No. 9.9;  
 Matches 77; Conservative 53; Mismatches 128; Indels 103; Gaps 20;  
 QY 276 GSGUGK--DYIQKSLQPIRISRSASLVGPNIGRPRKQVIALDDFMRAAGSYDVAW 333  
 DB 68 GSEWGFADACMKQ---YRLNKR--GPAVQAPRIQADKFL--YAGKVTYLETQ 118  
 QY 334 QHCYIDRVAWVDFLTKRLDLSQDIRKIQVNTYTPGKKIMLEGVTTSAAGTNL 393  
 DB 119 LHYQD---TVDVCSNTTDAGVAYGAHAHV--ARGRISARAVLTG----- 166  
 QY 394 SDVYAGFML-----NLGLANQGDVVRHSFPDGG-----YHLYDQNF 436  
 DB 167 --FMEBRVITGEYAEGRGSHABGLAARKKGFQNGRIKGTGPVRLKSYD--- 221  
 QY 437 NPLPDVWLSLYR---LIGPKVLAHVAGLQRRKPRGVRIDKLRIYAHCTNNHANY 492  
 DB 222 -----LSVMEKQADAIWRPSPFA-HV-----EINRPHADCIYNTNETHQD 263  
 QY 493 VRGSITLFLTNHRS---KKIKLAGT-----LRDGLVQYIOLPPGCGBGLKS 537  
 DB 264 IRE-----NFRSPFSGRIKAVGTRVCPSTIEDKVRKFPFDIRIQLYTER---EGD 313  
 QY 538 KSVQLANGOPLVWVD-----GTLPEIK---PRPRAGTAVIPVTVAGFYV-VGVNVA 586  
 DB 314 EELYINGLSCLEPDIDQEMIRITPGMERAVITRPAVADVAVLFPVQGIQLQTRVSG 373  
 QY 587 L 587

DB 374 L 374  
 RESULT 12  
 ID SUT2\_HUMAN STANDARD; PRT; 870 AA.  
 AC Q81WU5; Q96SG2; Q9H1H0; Q9UUR3; Q9ULH3;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Extracellular sulfatase sulf-2 precursor (EC 3.1.6.-) (HSulf-2).  
 GN SUT2 OR KIAA1247.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., CHARACTERIZATION, AND MUTAGENESIS OF  
 RP 88-CVS-CVS-89.  
 RC TISSUE=Lung;  
 RX MEDLINE=22370956; PubMed=12368295;  
 RA Morimoto-Tomita M., Uchimura K., Weib Z., Hemmerich S., Rosen S.D.,  
 RT "Cloning and characterization of two extracellular heparin-degrading  
 RT endosulfatases in mice and humans";  
 RL J. Biol. Chem. 277:49175-49185(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=20039619; PubMed=10574462;  
 RA Nagase T., Ishikawa K.-I., Kikuno R., Hirose M., Nomura N.,  
 RA Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XV.  
 RT for large proteins in vitro";  
 RL DNA Res. 6:337-345(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida J.P., Babage A.K., Baguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.K., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman U.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Coley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dham P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Grahame D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Lehaeaelaho M.H., Leverisa M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConachie L.J., McKay K., McMurray A.A.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
 RA Rice C.W., Ross M.T., Scott C.B., Sehra H.K., Showstreen R., Sims S.,  
 RA Skuce C.D., Smith M.L., Sodeitund C., Steward C.A., Sulston U.B.,  
 RA Swann R.M., Symmore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.W.,  
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20";  
 RL Nature 414:865-871(2001).  
 RN [4]  
 RP SEQUENCE OF 566-870 FROM N.A.  
 RC TISSUE=Colon;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schlier G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.T., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stalderon M., Soares M.B., Ronaldo M.P., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Rana S.S., Lounellano N.A., Peters G.J., Abrahamson R.D., Mullany S.J.,  
RA Bosak S.A., Mowbray P.J., McKernan K.U., Malek U.A., Gunaratne P.H.,  
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hultx S.W.,  
RA Villalón D.K., Murthy D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy U., Helton E., Ketterman M., Madan A.C., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.M., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz U., Myers R.M.,  
RA Butterfield A.C., Krzywinski M.I., Skalska U., Smalins D.E.,  
RA Schermer A., Schein J.E., Jones S.J.M., Marra M.A.,  
RA "Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences.";  
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -1- FUNCTION: Exhibits arylsulphatase activity and highly specific  
CC endoglucosaminase-6-sulphatase activity. It can remove sulfate from  
CC the C-6 position of glucosamine within specific subregions of  
CC intact heparin.  
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum and Golgi stack.  
CC -1- Also localized on the cell surface (By similarity).  
CC -1- TISSUE SPECIFICITY: Expressed at highest levels in the ovary,  
CC skeletal muscle, stomach, brain, uterus, heart, kidney and  
CC placenta.  
CC -1- MISCELLANEOUS: Shows maximal activity at pH 7.0 and 8.0.  
CC -1- SIMILARITY: Belongs to the sulfatase family.  
CC -1- CAUTION: Ref.3 (CAC17694) sequence differs from that shown due to  
CC erroneous gene model prediction.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and that statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; ABY01176; AAM76861.1; -; ALT INIT.  
CC EMBL; AB033073; BAA86561.2; -; ALT INIT.  
CC EMBL; AL354813; CAC39100.1; -; ALT INIT.  
CC EMBL; AL133001; CAB61349.1; -; ALT INIT.  
CC EMBL; AL034418; CAC17694.1; -; ALT SEQ.  
CC EMBL; BC020962; AAM20962.1; -; ALT INIT.  
CC InterPro: IPR000917, Sulfatase.  
CC Pfam: PF00884, Sulfatase; 1.  
CC PROSITE: PS00523; SULFATASE\_1; 1.  
CC PROSITE: PS00149; SULFATASE\_2; FALSE NEG.  
CC HydroLase; Signal; Glycoprotein; Endoplasmic reticulum; Golgi stack.  
CC FT SIGNAL 1 24  
CC FT CHAIN 25 870  
CC FT MOD\_RES 88 88  
CC FT  
CC CARBOHYD 65 65  
CC CARBOHYD 112 112  
CC CARBOHYD 132 132  
CC CARBOHYD 149 149  
CC CARBOHYD 171 171  
CC CARBOHYD 198 198  
CC CARBOHYD 241 241  
CC CARBOHYD 561 561  
CC CARBOHYD 608 608  
CC CARBOHYD 717 717  
CC CARBOHYD 754 754  
CC CARBOHYD 764 764  
CC COMPLET 88 89  
CC CONFICT 76 76  
CC SOURCE 870 AA; 100454 MW; 741069CE2774D73 CR664;  
CC  
CC Query Match 3.3%; Score 100.5; DB 1; Length 870;  
CC Best Local Similarity 19.3%; Pred. No. 16;  
CC Matches 65; Conservative 50; Mismatches 100; Indels 121; Gaps 18;

QY 86 SIHGWLDLSSKRLVTLARGLSPALFRGGRTPDLOFONLRNPAKSGGPGDYUK 145  
DB 151 SYVPGKKEV-----GLL-----KNSPFYVTLCRNGVKEHGSD---YAK 189  
QY 146 NYEDDIYRSVVALDKQCKCIAHPDVMVLQR-----EKAQMLVLLKPFSTNYEN 199  
DB 190 DYLLDLDITNDVSFFRFSKKYVHRPLVWISHAHPGPDSDA-----POYSLFPN 241  
QY 200 LILFARSLDLYNFDGSHLIFALNALRRNPNNSNSLSALILKYSASK-----252  
DB 242 ---ASGHITSYVAP-----NDKGM-----IRYIGPMPIHMEFT 276  
QY 253  
DB 277 NMLQRKRLQTLMSVDSMETIYMLVETGELDNTYIVY--TADHGHIQO-FGLVYKSKW 333  
QY 292 PIRIYRSALSY--GPNL--GRPRKNVYALD---GPMKVGSTVVDVWTHCYIDRKYV 344  
DB 334 PYEDDIVPYVYKQPNVAGCLAPHYVNLIDLAFTIIDIGLDIPA-----DMDKST- 386  
QY 345 VMDFLKRLDLYNFDGSHLIFALNALRRNPNNSNSLSALILKYSASK--IW 378  
DB 387 -----LKLDDT-----ERPYNREHLKKQKRWV 408  

RESULT 13  
MSH4 YEAST STANDARD; PRT; 878 AA.  
AC P40965;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE MSH4 OR YFL003C.  
GN MSH4 OR YFL003C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95094270; PubMed=8001134;  
RA Rose-Macdonald P., Roeder G.S.;  
RT "Mutation of a meiosis-specific Muts homolog decreases crossing over  
RT but not mismatch correction.";  
RL Cell 79:1069-1080(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RX MEDLINE=95400292; PubMed=7670463;  
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,  
RA Sasamura S.-I., Sasamura M., Tsuchiya Y., Soeda E., Yokoyama K.,  
RA Yamazaki M., Tashiro H., Eki T.;  
RT "Analysis of the nucleotide sequence of chromosome VI from  
RT Saccharomyces cerevisiae.";  
RL Nat. Genet. 10:261-268(1995).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX STRAIN=S288C / AB972;  
RX MEDLINE=96381249; PubMed=8789262;  
RA Naitou M., Ozawa M., Sasamura S.-I., Kobayashi M., Hagiwara H.,  
RA Shibata T., Hanaoka F., Watanabe K., Ono A., Yamazaki M., Tashiro H.,  
RA Eki T., Murakami Y.;  
RT "Sequencing of a 23 kb fragment from Saccharomyces cerevisiae  
RT chromosome VI.";  
RL Yeast 12:77-84(1996).  
CC -1- FUNCTION: Involved in meiotic recombination. Facilitate crossovers  
CC between homologs during meiosis.  
CC -1- SUBUNIT: Heterooligomer of MSH4 and MSH5.  
CC -1- SIMILARITY: Belongs to the DNA mismatch repair muts family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: U13999; AA65234.1; -  
 DR EMBL: D50617; BA09235.1; -  
 DR PIR: A55201; A55201.  
 DR GerMOnline; 140152; -  
 DR SGD; S0001891; MSH4.  
 DR GO; GO:0000228; C:nuclear chromosome, IDA.  
 DR GO; GO:0003677; F:DNA binding; IPI.  
 DR GO; GO:0007131; P:meiotic recombination; IMP.  
 DR InterPro; IPR000432; Muts\_C.  
 DR InterPro; IPR007860; Muts\_III.  
 DR InterPro; IPR007696; Muts\_III.  
 DR InterPro; IPR007861; Muts\_IV.  
 DR Pfam; PF05188; Muts\_II; 1.  
 DR Pfam; PF05192; Muts\_III; 1.  
 DR Pfam; PF05190; Muts\_IV; 1.  
 DR Pfam; PF00488; Muts\_V; 1.  
 DR ProDom; PD001263; Muts\_C; 1.  
 DR SMART; SM00534; Mutsac; 1.  
 DR SMART; SM00533; Mutsd; 1.  
 DR PROSITE; PS00466; DNA\_Mismatch\_REPAIR\_2; 1.  
 DR Meiosis; ATP-binding; DNA-binding.  
 DR NP\_BIND; 634 641 ATP (POTENTIAL).  
 DR CONFLICT; 811 812 MD -> IH (IN REF. 2 AND 3).  
 FT SEQUENCE 878 AA; 99220 MW; 0FALCI29CC3DBCA CRC64;

Query Match 3.3%; Score 100.5; DB 1; Length 878;  
 Best Local Similarity 19.3%; Pred. No. 16;  
 Matches 87; Conservative 72; Mismatches 182; Indels 109; Gaps 19;

QY 64 DVGKTPVRYVNER-----FLSLQDPSIIHGMFLPSKRLVTLARGISPAFLRF 115  
 DB 115 DISTRIGLCINCTGMYLSDPMDSQIYRYVHK--LQYOPTETILIPSSSLAPTV-- 169  
 QY 116 GGRKTDLPQNLNPAKSGRGGPDYVLYKNVEDDIVRSDVAL--DKQCKIAQPDVY 173  
 DB 170 -SKLATYIKK-NAEYTKIKESGRKC--NSQDGLAATIKYIMDTKDKLKEEIIDKT 224  
 QY 174 LVQREKAQMHV-----LIKEQFSNTYSLII--TAASLDKLYNFADG 216  
 DB 225 FALCAASAAISYEBEIIKSSRMVNAFRKRIQFEGEENTMLDSKTVAGLEIYEKDK 284  
 QY 217 SGLHLFALNALRRPNVNSSSALSILKYSKXKYNISWEL--GNENPVYRTWNGRA 273  
 DB 285 NGISL-----WKFLDTSTKMGQSLNSIQLPTDGSIEMLLEAEK 330  
 QY 274 VNGSQGKDYIQLKSLQPIRIYSRASLYGNITGRP--RKNVIALLDGEMKVAAGTVDAV 331  
 DB 331 ANDDLQKRLKEKMSFDDKLFSLICINHSAIKPDQRIYVLLIKETLQSVKSLKDAL 390  
 QY 332 TWQHCYIDGVVVMFLKTRLDL-----SPQIKIQVNVVTPGKKIMEGVVTS 386  
 DB 391 NDQ-----LIQSLISSTKTFNNDAIMEIEKLINSCINEDCVASSAIQLI 437  
 QY 387 AGGTNNLSDSYAGFLMNTLGMV--ANQGDIVVIRHSFF-----DHGYN-- 429  
 DB 438 N-----QRSYA--VKSDBNGLLDVSROIYKEVKEEFREVEDLTAKKIKINDHYDGA 488  
 QY 430 -----HLYQNFN-----PLPDYMLSLLYKR 450  
 DB 489 RGFYLRKROEFTDVAATLPDVFISRTIK 518

RESULT 14  
 LINI\_HUMAN STANDARD; PRT; 1259 AA.  
 ID LINI\_HUMAN  
 AC P08547;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 01-AUG-1988 (Rel. 08, Last sequence update)

DT 01-AUG-1988 (Rel. 08, Last annotation update)  
 DE LINE-1 reverse transcriptase homolog.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCB1\_TaxID=9606;  
 RN [1]  
 RP MEDLINE=86230917; PubMed=2423883;  
 RX Hattori M., Kubara S., Takenaka O., Sakaki Y.;  
 RA "L1 family of repetitive DNA sequences in primates may be derived  
 RT from a sequence encoding a reverse transcriptase-related protein";  
 RL Nature 321:625-628(1996).  
 CC -1- WISCCELAMZOUS: This sequence was constructed from an alignment of  
 CC published and unpublished sequences, determined in various  
 CC laboratories, belonging to the LINE-1 family.

DR PIR; A25313; GNHUL1.  
 DR InterPro; IPR005135; Exo\_endo\_phos.  
 DR InterPro; IPR000477; RYase.  
 DR Pfam; PF03372; Exo\_endo\_phos; 1.  
 DR Pfam; PF00078; rvt; 1.  
 DR RNA-directed DNA polymerase.  
 SQ SEQUENCE 1259 AA; 147112 MW; F3BC18A3803919F5 CRC64;

Query Match 3.3%; Score 100.5; DB 1; Length 1259;  
 Best Local Similarity 20.0%; Pred. No. 26;  
 Matches 131; Conservative 89; Mismatches 217; Indels 217; Gaps 35;

QY 50 DRAAGLKEKTLIILDVSTKQNPRT-----VNEFLSLQDPSIIHGMFLPSKRL 101  
 DB 130 DLORDLDSHTIMGDFYNT--FLSTLDSTRQKINKDQELN--SALHQA-----DL 176  
 QY 102 VTLARGLSPARLRGGKRTDLPQNLNPAKSGRGGPDYVLYKNVEDDIVRSDVALDKQ 161  
 DB 177 IDIYRLHP-----KSTETTFP-----SAPHHTYKTDHILGSKTILSKC 216  
 QY 162 KGCKI-----AQHEDVNLVQREKAQMHVTLK----- 190  
 DB 217 KRTETITNCSDHSAIKLEIRIKLTQNHSTWTKLNLINDVYVHMEKAETIKFETN 276  
 QY 191 EGFNTYNSNLITPARSDKLYNPADCGGHLIFPLNLLRRPNNSNNSSSALSILKYSAS 250  
 DB 277 ENKDTYQNLMDTKAKV-----CRGKFL--ALNHKKRQERS-KITDLISQLK--- 321  
 QY 251 KKYNISWELAGNEPNVNTWNGRAVNGSQLGKDYIQLKSL-----LQPIRIYSRASLYGP 304  
 DB 322 -----ELKQQTMSKASRQEIIRKIRAELEIETQTLQKIN--ESRSWPEKX 369  
 QY 305 -NIGRPKNVIA-----LDDGEMKVAAG--TVDAVT-----WQHCYIDG-RVYVMD 347  
 DB 370 NKIDRPLARLKKRERKNQIDITINDRGDITTPTEIQTTIREYKHLVANKLENEEMD 429  
 QY 348 -FLKTRLDLTDSD-----IRKIQVNVVTPGKKIMEGVVTSAGGTNNLSDS 386  
 DB 430 KFLDTYLPRLNGEVEASINRPISSSEIHALINS-LPNKK-----SPGEGGFLAEF 479  
 QY 397 Y-----AAGFLWNTGMLANQGDIVVIRHSFFD-----HGYNHLYDQNFNPDPY 442  
 DB 480 YQRYKEELVPEFL-KLPOSIEKEGI--LPNSFEASIIILIPRGRTTKEKFRPIS-- 533  
 QY 443 WLSLLYKRLKIPKYL-----VHVAGLQKRPDGV-----IRDLRIYACCTNNHNN 490  
 DB 534 -LNNIDAKILN-KILANQIQOHIKKLTHDOVGFIPLMOMGFWNIRKSNITIQHINRKOT 591  
 QY 491 NYVSGITLFTINLHRSRKIKYLAGTLRDLVHQLLYQPYGQGL-----K 536  
 DB 592 NHM-----IISD-----AEKAFDKIQQPFMLKPLNKIGIDTYLKIIRAIYDKP 636  
 QY 537 SKSVQNGQPLVWWDGTLPELKRPLRAGRTLVIPVTVGFWYKVVNALACR 590  
 DB 637 TANIILNQ-----KLEAPLKTGTGCGPLSPILNPVILEVLAIR 679



```

Db      2623  SS-----LIFLKPPLPGTYOFSLOIKSQSDSKIR-----SACHVWTVIP-----PTNLTW 2669
QY      237   NSSSALSLLKYASAKKNISWELGNEBPNNYRTMHGRAVN-----GSOLGXDIQPK 287
Db      2670  NIPSVI-----FATRNNTI-----PNLFHLPBGYSLSDDORTFSLISGSGKNSKLS 2717
QY      288   SLQPIRIYGRASLYGNIGRPRKNVIA--LIDGFMKVAGSTVDAYTMOHCYIDGRV-- 343
Db      2718  SGVYQVNV-----VGKDEKKEIVRILD-----DVADDTSKD--IEYHVSS 2758
QY      344   -----KXMDFLKTRLDLTLSDOIRKIQK-----VNTYTPGKKIMLEGVYVTSAGG 389
Db      2759  TLSNLKIPPIIDVECPRETEENLEYEITKOCRLFNSDVINTTIF-----VVTSPANS 2810
QY      390   TNNL 393
Db      2811  TNNL 2814

```

Search completed: May 6, 2004, 13:44:42  
 Job time : 20 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2004, 13:41:41 ; Search time 45 Seconds  
(without alignments)  
4150.817 Million cell updates/sec

Title: US-10-088-676-2

Perfect score: 3088  
Sequence: 1 MEVLCAFPAMPSSNSRPPA.....PYTMGFYVKNVNALACRFR 592

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

```
SPTREMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3078	99.7	592	4 Q9HB37	Q9HB37 homo sapien
2	3070	99.4	592	4 Q8WQ2	Q8WQ2 homo sapien
3	2785	90.2	548	4 Q8WQ1	Q8WQ1 homo sapien
4	2736	88.6	534	4 Q9HB38	Q9HB38 homo sapien
5	2446	79.2	480	4 Q9HB39	Q9HB39 mus sapien
6	1655	37.7	535	11 Q8XK3	Q8XK3 mus musculu
7	1165	37.7	536	11 Q9QZF8	Q9QZF8 rattus norv
8	1149	37.2	545	4 Q9UD39	Q9UD39 homo sapien
9	1144.5	37.1	543	4 Q9Y251	Q9Y251 homo sapien
10	1138	36.9	545	6 Q9MYX0	Q9MYX0 bos taurus
11	1033.5	33.5	523	13 Q90YX5	Q90YX5 gallus gall
12	625	20.2	515	5 Q8T108	Q8T108 bombyx mori
13	384	12.4	544	10 Q8H615	Q8H615 cryza sativ
14	382	12.4	521	10 Q9SDA1	Q9SDA1 arabidopsis
15	382	12.4	543	10 Q9F10	Q9F10 arabidopsis
16	334.5	10.8	516	10 Q9FLK8	Q9FLK8 arabidopsis

17	334.5	10.8	539	10 Q8L608	Q8L608 arabidopsis
18	314	10.2	536	10 Q9FZP1	Q9FZP1 arabidopsis
19	307.5	10.0	527	10 Q9LRC8	Q9LRC8 scutellaria
20	301.5	9.8	559	16 Q89FP9	Q89FP9 bradyrhizob
21	171	5.5	935	5 Q9VZ79	Q9VZ79 drosophila
22	155	5.0	190	10 Q8Z604	Q8Z604 arabidopsis
23	146	4.7	1260	5 Q15639	Q15639 dictyostell
24	128	4.1	1027	16 Q8Z028	Q8Z028 anabaena sp
25	127.5	4.1	1234	5 Q9Y1H7	Q9Y1H7 dictyostell
26	119	3.9	370	17 Q8TPM9	Q8TPM9 methanosarc
27	118.5	3.8	301	17 Q8TPM8	Q8TPM8 methanosarc
28	118.5	3.8	901	10 Q84ZU8	Q84ZU8 glycine max
29	115	3.8	459	10 Q9FVK3	Q9FVK3 glycine max
30	115	3.7	897	10 Q84ZV8	Q84ZV8 glycine max
31	114.5	3.7	3218	5 Q8ILX0	Q8ILX0 plasmodium
32	114	3.7	424	5 Q8T8M7	Q8T8M7 caenorhabdi
33	114	3.7	556	12 Q805J2	Q805J2 strawberry
34	110.5	3.6	1471	5 Q9VU00	Q9VU00 drosophila
35	110	3.6	174	10 Q9ATW5	Q9ATW5 zea mays (m
36	110	3.6	370	17 Q8TR29	Q8TR29 methanosarc
37	110	3.6	370	17 Q8TR26	Q8TR26 methanosarc
38	109	3.5	3317	16 Q8EWP8	Q8EWP8 mycoplasma
39	108.5	3.5	1156	16 Q8KFM2	Q8KFM2 chlorobium
40	107.5	3.5	1053	2 P71329	P71329 fibrobacter
41	107	3.5	895	10 Q84RV3	Q84RV3 glycine max
42	107	3.5	915	16 Q7U2L6	Q7U2L6 prochloroco
43	106	3.4	556	12 Q80HR4	Q80HR4 strawberry
44	106	3.4	565	10 Q9L748	Q9L748 arabidopsis
45	106	3.4	762	5 Q27701	Q27701 macrobdeilia

#### ALIGNMENTS

##### RESULT 1

Q9HB37 PRELIMINARY: PRT, 592 AA.

```
AC Q9HB37, 16, Created
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Heparanase-like protein HPA2C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20483645; PubMed=11027606;
RA McKenzie E., Tyson K., Stamps A., Smith P., Turner P., Barry R.,
RA Hirock M., Patel S., Barry E., Stuberfield C., Terrett J., Page M.;
RT "Cloning and Expression Profiling of Hpa2, a Novel Mammalian.
RT Heparanase Family Member."
RL Biochem. Biophys. Res. Commun. 276:1170-1177(2000).
RL EMBL; AF282887; AAG23423.1; -.
DR GO; GO:0005622; C:intracellular; TAS.
DR GO; GO:0030305; F:heparanase activity; TAS.
DR InterPro; IPR005199; Glyco_hydro_79N.
DR Pfam; PF03662; Glyco_hydro_79N; 1.
SQ SEQUENCE 592 AA; 66580 MW; 95C384AD9A74258E CRC64;
```

Query Match 99.7%; Score 3078; DB 4; Length 592;  
Best Local Similarity 99.7%; Pred. No. 1.8e-235;  
Matches 590; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MEVLCAFPAMPSSNSRPPACIAPGALYIALIHLISQAGDRRLPVDRRAAGKENTL 60
DB 1 MEVLCAFPAMPSSNSRPPACIAPGALYIALIHLISQAGDRRLPVDRRAAGKENTL 60
QY 61 ILLDVSTKNPVATVENFLSIQLDPSIIHDGMDLFLSSKRLVTTARGSPAPFRGGKRT 120
DB 61 ILLDVSTKNPVATVENFLSIQLDPSIIHDGMDLFLSSKRLVTTARGSPAPFRGGKRT 120
```



```

QY 121 DFLQFQNLNRPAPKSRGGPGPDYLLKNYEDDIVRSVDVALDKKQCKIAQHPDVMVLQREK 180
DB 121 DFLQFQNLNRPAPKSRGGPGPDYLLKNYEDDIVRSVDVALDKKQCKIAQHPDVMVLQREK 180
QY 181 AAQOMHVLVLEQFSNTYSNLILTARSIDLKYNFADCSGHLIFALNALRRPNNSWNSS 240
DB 181 AAQOMHVLVLEQFSNTYSNLILTARSIDLKYNFADCSGHLIFALNALRRPNNSWNSS 240
QY 241 ALSLKYSASKKNISWELGNEBNRYRWGAVNSQGLKDYIQLKSLLOPIRISRAS 300
DB 241 ALSLKYSASKKNISWELGNEBNRYRWGAVNSQGLKDYIQLKSLLOPIRISRAS 300
QY 301 LYGPNIGRPKNVIALDGFPMKVASGVDAVWQHOCYIDGRVYKWDPLKTRLDLTLSQ 360
DB 301 LYGPNIGRPKNVIALDGFPMKVASGVDAVWQHOCYIDGRVYKWDPLKTRLDLTLSQ 360
QY 361 IRKIQKVVNTYTPGKKIWEQVVTTSAGGTNNLSDSYAAGFLMLNTLGMLANQIDVVR 420
DB 361 IRKIQKVVNTYTPGKKIWEQVVTTSAGGTNNLSDSYAAGFLMLNTLGMLANQIDVVR 420
QY 421 HSFPHGYNHLVDQNFNPLPDYMLSLLYKELIGPKYLAHVAGLQKRPGRVIRDKLRI 480
DB 421 HSFPHGYNHLVDQNFNPLPDYMLSLLYKELIGPKYLAHVAGLQKRPGRVIRDKLRI 480
QY 481 YACHTNHNHNYVGSITLFTINLHRSRKKIKLAGTLRDKLVHGYLLQPYGQGLSKSV 540
DB 481 YACHTNHNHNYVGSITLFTINLHRSRKKIKLAGTLRDKLVHGYLLQPYGQGLSKSV 540
QY 541 QLNQGPLVWDDGTLPELKRPLRAGRTLVIPYTMGFVYKVNALACRYR 592
DB 541 QLNQGPLVWDDGTLPELKRPLRAGRTLVIPYTMGFVYKVNALACRYR 592

```

## RESULT 2

```

Q88WQ2 ID Q88WQ2 PRELIMINARY; PRT; 592 AA.
AC 088WQ2;
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Heparanase 2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Pessague Safontas B.J.O.P.S.;
RU Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Legoux P., Legoux R., O'Brien D., Salome M.;
RU Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ299719; CAC82491.1;
DR Genev; HGNC:18374; HPS2.
DR InterPro; IPR005199; Glyco_hydro_79N.
DR Pfam; PF03662; Glyco_hydro_79N; I.
SQ SEQUENCE 592 AA; 66520 MW; 9478841FEACD558B CRC64;

Query Match 99.4%; Score 3070; DB 4; Length 592;
Best Local Similarity 99.5%; Pred. No. 7.9e-235;
Matches 589; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 121 DFLQFQNLNRPAPKSRGGPGPDYLLKNYEDDIVRSVDVALDKKQCKIAQHPDVMVLQREK 180
DB 121 DFLQFQNLNRPAPKSRGGPGPDYLLKNYEDDIVRSVDVALDKKQCKIAQHPDVMVLQREK 180
QY 181 AAQOMHVLVLEQFSNTYSNLILTARSIDLKYNFADCSGHLIFALNALRRPNNSWNSS 240
DB 181 AAQOMHVLVLEQFSNTYSNLILTARSIDLKYNFADCSGHLIFALNALRRPNNSWNSS 240
QY 241 ALSLKYSASKKNISWELGNEBNRYRWGAVNSQGLKDYIQLKSLLOPIRISRAS 300
DB 241 ALSLKYSASKKNISWELGNEBNRYRWGAVNSQGLKDYIQLKSLLOPIRISRAS 300
QY 301 LYGPNIGRPKNVIALDGFPMKVASGVDAVWQHOCYIDGRVYKWDPLKTRLDLTLSQ 360
DB 301 LYGPNIGRPKNVIALDGFPMKVASGVDAVWQHOCYIDGRVYKWDPLKTRLDLTLSQ 360
QY 361 IRKIQKVVNTYTPGKKIWEQVVTTSAGGTNNLSDSYAAGFLMLNTLGMLANQIDVVR 420
DB 361 IRKIQKVVNTYTPGKKIWEQVVTTSAGGTNNLSDSYAAGFLMLNTLGMLANQIDVVR 420
QY 421 HSFPHGYNHLVDQNFNPLPDYMLSLLYKELIGPKYLAHVAGLQKRPGRVIRDKLRI 480
DB 421 HSFPHGYNHLVDQNFNPLPDYMLSLLYKELIGPKYLAHVAGLQKRPGRVIRDKLRI 480
QY 481 YACHTNHNHNYVGSITLFTINLHRSRKKIKLAGTLRDKLVHGYLLQPYGQGLSKSV 540
DB 481 YACHTNHNHNYVGSITLFTINLHRSRKKIKLAGTLRDKLVHGYLLQPYGQGLSKSV 540
QY 541 QLNQGPLVWDDGTLPELKRPLRAGRTLVIPYTMGFVYKVNALACRYR 592
DB 541 QLNQGPLVWDDGTLPELKRPLRAGRTLVIPYTMGFVYKVNALACRYR 592

```

## RESULT 3

```

Q88WQ1 ID Q88WQ1 PRELIMINARY; PRT; 548 AA.
AC 088WQ1;
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 24, Last annotation update)
DE Heparanase 3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Pessague Safontas B.J.O.P.S.;
RU Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Legoux P., Legoux R., O'Brien D., Salome M.;
RU Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ299720; CAC82492.1;
DR InterPro; IPR005199; Glyco_hydro_79N.
DR Pfam; PF03662; Glyco_hydro_79N; I.
SQ SEQUENCE 548 AA; 61771 MW; B8986303RC73A60A CRC64;

Query Match 90.2%; Score 2785; DB 4; Length 548;
Best Local Similarity 99.3%; Pred. No. 2.9e-212;
Matches 535; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

Db      121 DFLQFONLRNPAKSRGGPGPDYLLKYNEDDIIVRSVALDKOKCKIAQHPDVMLIQREK 180
Qy      181 AAOQMHVLLKQGSNTYSNLTARSLDKYNFADCSGLHIFALNALRRNPNNSSS 240
Db      181 AAOQMHVLLKQGSNTYSNLTARSLDKYNFADCSGLHIFALNALRRNPNNSSS 240
Qy      241 ALSLKYSASKKYNISWELGNEPNNTYTMGRAVNGSOLGKDYIOLKSLQPIRISRAS 300
Db      241 ALSLKYSASKKYNISWELGNEPNNTYTMGRAVNGSOLGKDYIOLKSLQPIRISRAS 300
Qy      301 LYPENIGRPKNVIALLDGFMKAVGSTVDATWQHCHTIDGRVYKVMDFLKTLLDITLSQ 360
Db      301 LYPENIGRPKNVIALLDGFMKAVGSTVDATWQHCHTIDGRVYKVMDFLKTLLDITLSQ 360
Qy      421 IRKIQKVNNTYTPGKKIMLEGVVTTSAGTNNLSDSYAAGFLMNTLGMLANOGIDVIR 420
Db      421 IRKIQKVNNTYTPGKKIMLEGVVTTSAGTNNLSDSYAAGFLMNTLGMLANOGIDVIR 420
Qy      481 YACHTNHNHYVNGSITLFTIINLHRSRKKIKLAGTRDKLVHGYLLQPIGQEGLSKSV 540
Db      481 YACHTNHNHYVNGSITLFTIINLHRSRKKIKLAGTRDKLVHGYLLQPIGQEGLSKSV 540

```

## RESULT 4

```

ID      Q9HB38      PRELIMINARY;      PRT;      534 AA.
AC      Q9HB38;
DT      01-MAR-2001 (TREMBlrel. 16, Created)
DT      01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT      01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE      Heparanase-like protein HPa2b.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      MEDLINE=20483645; PubMed=11027606;
RA      McKenzie E., Tyson K., Stamps A., Smith P., Turner P., Barry R.,
RT      "Cloning and Expression Profiling of Hpa2, a Novel Mammalian
RT      Heparanase Family Member."
RL      Biochem. Biophys. Res. Commun. 276:1170-1177(2000).
DR      EMBL; AF282886; AAG23422.1;
DR      InterPro; IPR005199; Glyco_hydro_79N.
DR      Pfam; PF03662; Glyco_hydro_79n; 1.
SQ      SEQUENCE 534 AA; 60063 MW; C3DESE900CB38C4 CRC64;

```

Query Match 88.6%; Score 2736; DB 4; Length 534;  
 Best Local Similarity 89.9%; Pred. No. 2.2e-208; Indels 58; Gaps 1;  
 Matches 532; Conservative 1; Mismatches 1;

```

Qy      1 MRYLCAPFEAMPSSNSRPPACIAPGALYIALILHLISSSQAGDRRLPVDRAGLKEKTL 60
Db      1 MRYLCAPFEAMPSSNSRPPACIAPGALYIALILHLISSSQAGDRRLPVDRAGLKEKTL 60
Qy      61 ILDDVSTKNPVRTVNTENFLSLQDPSIIHGMDFLSSKRLVTLARGLSPAFIRFGGKRT 120
Db      61 ILDDVSTKNPVRTVNTENFLSLQDPSIIHGMDFLSSKRLVTLARGLSPAFIRFGGKRT 120
Qy      121 DFLQFONLRNPAKSRGGPGPDYLLKYNEDDIIVRSVALDKOKCKIAQHPDVMLIQREK 180
Db      121 DFLQFONLRNPAKSRGGPGPDYLLKYNEDDIIVRSVALDKOKCKIAQHPDVMLIQREK 180
Qy      181 AAOQMHVLLKQGSNTYSNLTARSLDKYNFADCSGLHIFALNALRRNPNNSSS 240
Db      181 AAOQMHVLLKQGSNTYSNLTARSLDKYNFADCSGLHIFALNALRRNPNNSSS 240

```

```

Qy      241 ALSLKYSASKKYNISWELGNEPNNTYTMGRAVNGSOLGKDYIOLKSLQPIRISRAS 300
Db      241 -----EPNNYRTMGRAVNGSOLGKDYIOLKSLQPIRISRAS 242
Qy      301 LYPENIGRPKNVIALLDGFMKAVGSTVDATWQHCHTIDGRVYKVMDFLKTLLDITLSQ 360
Db      301 LYPENIGRPKNVIALLDGFMKAVGSTVDATWQHCHTIDGRVYKVMDFLKTLLDITLSQ 360
Qy      421 IRKIQKVNNTYTPGKKIMLEGVVTTSAGTNNLSDSYAAGFLMNTLGMLANOGIDVIR 420
Db      421 IRKIQKVNNTYTPGKKIMLEGVVTTSAGTNNLSDSYAAGFLMNTLGMLANOGIDVIR 420
Qy      481 YACHTNHNHYVNGSITLFTIINLHRSRKKIKLAGTRDKLVHGYLLQPIGQEGLSKSV 540
Db      481 YACHTNHNHYVNGSITLFTIINLHRSRKKIKLAGTRDKLVHGYLLQPIGQEGLSKSV 540
Qy      541 QLNQGPLVMVDGTLPELKPRLRAGRTLVIPVTMGFTYKVNVALACRYR 592
Db      541 QLNQGPLVMVDGTLPELKPRLRAGRTLVIPVTMGFTYKVNVALACRYR 592

```

## RESULT 5

```

ID      Q9HB39      PRELIMINARY;      PRT;      480 AA.
AC      Q9HB39;
DT      01-MAR-2001 (TREMBlrel. 16, Created)
DT      01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT      01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE      Heparanase-like protein HPa2a.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      MEDLINE=20483645; PubMed=11027606;
RA      McKenzie E., Tyson K., Stamps A., Smith P., Turner P., Barry R.,
RT      "Cloning and Expression Profiling of Hpa2, a Novel Mammalian
RT      Heparanase Family Member."
RL      Biochem. Biophys. Res. Commun. 276:1170-1177(2000).
DR      EMBL; AF282886; AAG23421.1;
DR      PIR; JG7506; JG7506.
DR      InterPro; IPR005199; Glyco_hydro_79N.
DR      Pfam; PF03662; Glyco_hydro_79n; 1.
SQ      SEQUENCE 480 AA; 53900 MW; F75F89F67AC1FF83 CRC64;

```

Query Match 79.2%; Score 2446; DB 4; Length 480;  
 Best Local Similarity 80.9%; Pred. No. 1.9e-185; Indels 112; Gaps 1;  
 Matches 479; Conservative 1; Mismatches 0;

```

Qy      1 MRYLCAPFEAMPSSNSRPPACIAPGALYIALILHLISSSQAGDRRLPVDRAGLKEKTL 60
Db      1 MRYLCAPFEAMPSSNSRPPACIAPGALYIALILHLISSSQAGDRRLPVDRAGLKEKTL 60
Qy      61 ILDDVSTKNPVRTVNTENFLSLQDPSIIHGMDFLSSKRLVTLARGLSPAFIRFGGKRT 120
Db      61 ILDDVSTKNPVRTVNTENFLSLQDPSIIHGMDFLSSKRLVTLARGLSPAFIRFGGKRT 120
Qy      121 DFLQFONLRNPAKSRGGPGPDYLLKYNEDDIIVRSVALDKOKCKIAQHPDVMLIQREK 180
Db      121 DFLQFONLRNPAKSRGGPGPDYLLKYNEDDIIVRSVALDKOKCKIAQHPDVMLIQREK 180
Qy      181 AAOQMHVLLKQGSNTYSNLTARSLDKYNFADCSGLHIFALNALRRNPNNSSS 240
Db      181 AAOQMHVLLKQGSNTYSNLTARSLDKYNFADCSGLHIFALNALRRNPNNSSS 240
Qy      241 ALSLKYSASKKYNISWELGNEPNNTYTMGRAVNGSOLGKDYIOLKSLQPIRISRAS 300
Db      241 -----EPNNYRTMGRAVNGSOLGKDYIOLKSLQPIRISRAS 300

```

DB 150 -----EENRYTHGRAVNSQJGKDYIQKSLQPIRYSRA 188

QY 301 LYGPNIGRPRKNVIALDGFMRVAGSTVDVAVTWOHCYIDGRVYKMDPLKTLDTLSDQ 360

DB 189 LYGPNIGRPRKNVIALDGFMRVAGSTVDVAVTWOHCYIDGRVYKMDPLKTLDTLSDQ 248

QY 361 IRIQVNVNTPYPRGKIMLEGVVTTSAAGTNNLSDSYAGFLMNTLQMLANQSIDVYIR 420

DB 249 IRIQVNVNTPYPRGKIMLEGVVTTSAAGTNNLSDSYAGFLMNTLQMLANQSIDVYIR 308

QY 421 HSPFDGYNHLYVDQNFNPLPDYWLSTLYRKLIGPKYLAHVAGLQKPRPRGVRIRDKLRI 480

DB 309 HSPFDGYNHLYVDQNFNPLPDYWLSTLYRKLIGPKYLAHVAGLQKPRPRGVRIRDKLRI 368

QY 481 YACHTNHNHNYRGSITLFTIHLRSRKKIKLAGTLRDKLVHGYLLQPYGQGLKSKSV 540

DB 369 YACHTNHNHNYRGSITLFTIHLRSRKKIKLAGTLRDKLVHGYLLQPYGQGLKSKSV 428

QY 541 QLNGQPLVWDDGTLPELKRPLRAGRTLVIPVTMGFYVYKVNALACRYR 592

DB 429 QLNGQPLVWDDGTLPELKRPLRAGRTLVIPVTMGFYVYKVNALACRYR 480

## RESULT 6

Q8K3K3 PRELIMINARY; PRT; 535 AA.

ID Q8K3K3

AC Q8K3K3; PRELIMINARY; PRT; 535 AA.

DT 01-OCT-2002 (TREMBLrel. 22, Created)

DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE Heparanase.

GN HPSE OR HPA.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI\_Taxid=10950;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=PEB;

RA Mao H.-Q., Navarro E., Patel S., Sargent D., Koo H., Wan H., Plata A., Zhou Q., Ludwig D., Bohlen P., Kusie P.; "Cloning, expression, and purification of mouse heparanase."

RT Protein Expr. Purif. 0:0-0(2002).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CS7BL/6J; TISSUE=Thymus;

RX MEDLINE=2354683; PubMed=12466851;

RA The FANTOM Consortium,

RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."

RL Nature 420:563-573(2002).

DR EMBL; AY07467; AAL76083.1; -.

DR EMBL; AK040471; BAC30600.1; -.

DR MGD; MGI:1343124; Hpsa.

DR InterPro; IPR005199; Glyco\_hydro\_79N.

DR Pfam; PF03662; Glyco\_hydro\_79n; I.

SQ SEQUENCE 535 AA; 6065 MW; 6E73A8302FB8A0DF CRC64;

Query Match 37.7%; Score 1165.5; DB 11; Length 535;

Best Local Similarity 46.2%; Pred. No. 9.9e-84;

Matches 244; Conservative 81; Mismatches 176; Indels 27; Gaps 6;

QY 63 LDVSTNPRVTNENFLSLQDPSIHD-GWLDPLSKRLVTLARGSLPAFRFGKRTD 121

DB 33 LRYVTRPRSVSPSLSTIDSLATDPRFLTFLGSPRLARSLRAYIRFGGTKD 92

QY 122 FLOFQNLNPNKSGRGPDPYILKNTEDIIVSDVALDKQKCKIAQHPDVLVLOREKA 181

DB 93 FLTF-----DPDKPTSEERSYWSQVNHDIQSESPV-----SAAVLRKLQVEWP 137

QY 182 AQCHLVLLKEQFNTSNILTRSLDKLYNPDGSLHLIPALNALRPNNSWSSA 241

DB 138 FQ-ELLRLRQYQKERNSTYSRSSVDMLXSPAKSGDLIRGLNALRTPDLRMSSNA 196

QY 242 LSLIKYSASKKYNI SWELGNEPNNRYTHGRAVNSQJGKDYIQKSLQPIRYSRA 301

DB 197 QLLDYSKSSKYNI SWELGNEPNSFWKKAHLLIDGLOQEDFVELHKLQK-SAFONAKI 255

QY 302 YGNPGRPRKNVIALDGFMRVAGSTVDVAVTWOHCYIDGRVYKMDPLKTLDTLSDQ 361

DB 256 YGNPGRPRKNVIALDGFMRVAGSTVDVAVTWOHCYIDGRVYKMDPLKTLDTLSDQ 315

QY 362 RIKQVNVNTPYPRGKIMLEGVVTTSAAGTNNLSDSYAGFLMNTLQMLANQSIDVYIR 421

DB 316 RIKQVNVNTPYPRGKIMLEGVVTTSAAGTNNLSDSYAGFLMNTLQMLANQSIDVYIR 375

QY 422 SFPDGHNLVDQNFNPLPDYWLSTLYRKLIGPKYLAHVAGLQKPRPRGVRIRDKLRI 481

DB 376 SFPDGHNLVDQNFNPLPDYWLSTLYRKLIGPKYLAHVAGLQKPRPRGVRIRDKLRI 426

QY 482 AHCNTNHNHNYRGSITLFTIHLRSRKKIKLAGTLRDKLVHGYLLQPYGQGLKSKSV 541

DB 427 AHCNTNHNHNYRGSITLFTIHLRSRKKIKLAGTLRDKLVHGYLLQPYGQGLKSKSV 486

QY 542 LNGQPLVWDDGTLPELKRPLRAGRTLVIPVTMGFYVYKVNALACRYR 592

DB 487 LNGQPLVWDDGTLPELKRPLRAGRTLVIPVTMGFYVYKVNALACRYR 534

## RESULT 7

Q9QZF8 PRELIMINARY; PRT; 536 AA.

ID Q9QZF8

AC Q9QZF8; PRELIMINARY; PRT; 536 AA.

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Heparanase.

GN HEP.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

NCBI\_Taxid=10116;

RN [1]

RP SEQUENCE FROM N.A.

RA Podjma K.A., Yokote H., Sakaguchi K., Ikura M., Yanagishita M.; "Heparanase from parathyroid cell line."

RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF184967; AAF04563.1; -.

DR InterPro; IPR005199; Glyco\_hydro\_79N.

DR Pfam; PF03662; Glyco\_hydro\_79n; I.

SQ SEQUENCE 536 AA; 60569 MW; 6208B1FD9E28421 CRC64;

Query Match 37.7%; Score 1165; DB 11; Length 536;

Best Local Similarity 45.1%; Pred. No. 1.1e-83;

Matches 248; Conservative 87; Mismatches 187; Indels 28; Gaps 7;

QY 42 GDRRPLPVDPAAG-LKEKTLILDVSTKPNVRTNENFLSLQDPSIHD-GWLDPLSKR 99

DB 12 GRRLALQGRPAAGAPKVDVLEFYTKRFGQSVSPSLSTIDSLATDPRFLTFLGSP 71

QY 100 RLVTIARGSLPAFRFGKRTDPLQFQNLNPNKSGRGPDPYILKNTEDIIVSDVALD 159

DB 72 RLRLARGSLPAFRFGKRTDPLQFQNLNPNKSGRGPDPYILKNTEDIIVSDVALD 123

QY 160 KQKCKIAQHPDVLVLOREKAQMLVLLKEQFNTSNILTRSLDKLYNPDGSLHLIPAL 219

DB 124 -----RVSADVLRLKQEWMPQ-ELLRLRQYQKERNSTYSRSSVDMLXSPAKSGDL 175

QY 220 HLIPALNALRPNNSWSSA-SLTKYSASKKYNI SWELGNEPNNRYTHGRAVNSQJGK 279

DB 176 DLIFGIALALRTPDLRMSSNAQQLLNTYSSKYNISWELGNEPNSFWKKAQISDGLQ 235

QY 280 GKDYLQKSLQPIRYSRA-SLTKYSASKKYNI SWELGNEPNNRYTHGRAVNSQJGK 339

DB 236 GEDFVELHKLQK-SAQONAKI-LGPDIGQPRGVTVLNLSFLAAGEVIDSLTWHHYLN 294

QY 340 GRVVKWDFKTRLLDPLSDQIRKIQKVVNTYTPGKKIMLEGVTTSSAGTNNLSDSYAA 339  
 DB 295 GRVATKEDPLSSDVLDFIILSVQKILKATKEMTPGKKVIMLGESTSAVGGGAPILSTNFAA 354  
 QY 400 GFLMLNTLQMLANOGIDIVYHSHFFDHGYNHLVDQNNPPLPDVYLSTLYRLGPKTLAV 459  
 DB 355 GFWMLDLKGLSAQGLIEVVMQVFFGAGNVHLVDENEPPLPDVYLSLFFKLVGPKYLM 414  
 QY 460 HVAGLQKRRPRGRVIRPKLRIYACTNNHNNHYVRSITLFIINLSRKKIKLAGTLRD 519  
 DB 415 RVKRPD-----RSLKRYLAHCTVYHPRYHREGDLTYLVNLHVHTLKLPPMFS 465  
 QY 520 KLVHGYLQPYGGGGLSKSVQVNGQPLVWVDDGTLDELKRPPLRARTVYIPVTMGFY 579  
 DB 466 RPVVKYLLKPPGSGLSKSVQVNGQTLKVDQTLPALTEKPLPAGSSISVPAFSYGF 525  
 QY 580 VVKVYNALAC 589  
 DB 526 VIRNAKTIAC 535

## RESULT 8

QY 09UL39 PRELIMINARY; PRT; 545 AA.

AC 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Heparanase.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=2029546; PubMed=10764835;  
 RA Dempsey L.A., Plummer T.B., Coombes S.L., Platt J.L.;  
 RT "Heparanase expression in invasive trophoblasts and acute vascular  
 damage.";  
 RL Glycobiology 10:467-475 (2000).  
 DR EMBL; AF084467; AAS4516.1; -;  
 DR InterPro; IPR005199; Glyco\_hydro\_79N.  
 DR Pfam; PF03662; Glyco\_hydro\_79n; 1.  
 SQ SEQUENCE 545 AA; 61418 MW; 6780ACD73C5A9A1 CRC64;

Query Match 37.2%; Score 1149; DB 4; Length 545;

Best Local Similarity 43.7%; Pred. No. 2, 1e-82;

Matches 251; Conservative 82; Mismatches 206; Indels 36; Gaps 8;

QY 17 RPPACIAGALYIALLLHLSQAGRRPLPVRAAGLKEKTLILDVSTKNVRYTNE 76  
 DB 4 RSKRPLRPPLMLLLGLPLSPGALPRPAQ-----QDVVDLDFGQPLHLVSP 56  
 QY 77 NFLSLQDPLSIHD-GWLDPLSSKRLVTLARGLSAFPRFGCKTDPLOFQMLNPAKSR 135  
 DB 57 SFLSYTITANLATDRFLILGLSPRLTARGLSPAYIRFGSTKDFILF---DPKES 112  
 QY 136 GGPQGVYLYKVEDDIVSDVALDKQKCKIAQ-HPDVLYLOREKAQMLVLLKEQFS 194  
 DB 113 TFEBSYVQSQVNOI-----CKGSIPTDVEELRLRLEWYQEDL-LIREHYQ 159  
 QY 195 NTYSNLITANSGLDYLPNADCGSLHLFALNALRRNNNSWSSALSILKYSASKYN 254  
 DB 160 KKFKNSTYSRSSVDVLYTFANCOSGLDLFGNALRLRTDLOMNSSNALLLDYCSSKGYN 219  
 QY 255 ISMELGNPNRYRTWAGRVANGSOLGKYIOLKSLQPIRYSRAISVGNIGRPKKVY 314  
 DB 220 ISMELGNPNRYRTWAGRVANGSOLGKYIOLKSLQPIRYSRAISVGNIGRPKKVY 278  
 QY 315 ALIDGFMKVAAGSTVADVATVQHCYIDGRVVKWDFKTRLLDPLSDQIRKIQKVVNTYTPG 374

DB 279 KMLKSPFLKAGGEIVDSVTWHHYIANGRTATREDFLNPVLDFIISVQKQVVESTRP 338  
 QY 375 KXIMLEGVTTSSAGTNNLSDSYAGFLMLNTLQMLANOGIDIVYIRHSFDPHGNHLVDQ 434  
 DB 339 KKVWLGESTSAVGGGAPILSDTPAAGFWMLDKLGLSAMGLIEVVMQVFFGAGNVHLVDE 398  
 QY 435 NNPNPLPDVYLSLTYRLGPKTLAVHAGLQKRRPRGRVINDKLRIVACNNHNNHYR 494  
 DB 399 NPDPLPDVYLSLFFKLVGTKYLMAVQSSKRR-----KLRVLAHCTVTDNPRYKE 449  
 QY 495 GSITLFIINLSRKKIKLAGTLRDKLVHGYLQPYGGGGLSKSVQVNGQPLVWVDDGT 554  
 DB 450 GDLITLAIINLHVHTYKTLRLPYPSNKKQVDKILRPLGPHGLLSKSVQVNGTLKVVDDGT 509  
 QY 555 LPBLKRPPLRARTVYIPVTMGFYVGNVNALAC 569  
 DB 510 LPPLMEKPLRPGSLGLPFAFSYFFVIRNAKVAAC 544

## RESULT 9

QY 09Y251 PRELIMINARY; PRT; 543 AA.

AC 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE HEPARANASE (HPESE protein).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=99321249; PubMed=10395326;  
 RA Huliett M.D., Freeman C., Handorf B.J., Baker R.T., Harris M.J.,  
 RA Parish C.R.,  
 RT "Cloning of mammalian heparanase, an important enzyme in tumor  
 invasion and metastasis.";  
 RL Nat. Med. 5:803-809 (1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Vlodavetsky I., Friedman Y., Elkin M., Ainsorn H., Alzmon R.,  
 RA Izhai-Michaeli R., Bitan M., Pappo O., Perez T., Michal I.,  
 RA Spector L., Becker I.;  
 RT "Mammalian heparanase: a novel gene involved in tumor progression and  
 metastasis.";  
 RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99377052; PubMed=10446189;  
 RA Toyoshima M., Nakajima M.;  
 RT "Human heparanase. Purification, characterization, cloning, and  
 expression.";  
 RL J. Biol. Chem. 274:24153-24160 (1999).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=9935379; PubMed=10405343;  
 RA Kusile P.H., Hulmes J.D., Ludwig D., Patel S., Navarro E.C.,  
 RA Seddon A.P., Giorgio N.A., Bohlen P.;  
 RT "Cloning and Functional Expression of a Human Heparanase Gene.";  
 RL Biochem. Biophys. Res. Commun. 261:183-187 (1999).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Pancreas;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh P.,  
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carinici P., Prange C.,  
 RA Raba S.S., Loggellano N.A., Peters G.J., Abrahamson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Ginnatone P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.X., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Mizny D.M., Sodergren E.U., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez M.I., Skalska U., Schmutz J., Myers R.M., Buterfield Y.S.,  
 RA Jones S.J., Maier M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [6].

RP SEQUENCE FROM N.A.  
 RC Tissue=Pancreas;  
 RA Strubberg R.;  
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF165154; AAD45379.1; -  
 DR EMBL; AF144325; AAD4342.1; -  
 DR EMBL; AF155510; AAD54941.1; -  
 DR EMBL; AF152376; AAD4669.1; -  
 DR EMBL; BC051321; AAH51321.1; -  
 DR Gene; HGNC:5164; HPSB.  
 DR GO; GO:0004566; F-beta-glucuronidase activity; TAS.  
 DR GO; GO:0007125; P:invasive growth; TAS.  
 DR GO; GO:0006029; P:proteoglycan metabolism; TAS.  
 DR InterPro; IPR005199; Glyco\_hydro\_79N.  
 DR Pfam; PF03662; Glyco\_hydro\_79n; I.  
 SQ SEQUENCE 543 AA; 61176 MW; AD262EC267334AB2 CRC64;

Query Match 37.1%; Score 1144.5; DB 4; Length 543;  
 Best Local Similarity 43.2%; Pred. No. 4, 7e-82;  
 Matches 248; Conservative 83; Mismatches 190; Indels 53; Gaps 9;

QY 18 PACACGATLALHLHSLSSQAGRRPLPVDRAGLKEKTLILDVSTKNPRTVEN 77  
 DB 20 PLAGSPGAL-----PRPA-----QAGVVDLDFTFQEPHLSPS 55  
 QY 78 FLSGLDPSIIHD-GWLDPLSKRLVTLARGLSPALFPGGKRTDPLQFONLNPARG 136  
 DB 56 FLSVITDANLADTPRFLILGSPKRLTLARGLSPALFPGGKRTDPLF-----DKKEST 111  
 QY 137 GGPDPYLLKNVEDDIVRSVALDKQCKCIAQ-HEDVNLVLOREKAAQMHVLVLEKESN 195  
 DB 112 FEERSYWGQVNOI-----CKYGSIPDVEEKLLEWFOEOL-LNREHYQK 158  
 QY 196 TYSNLTILTRSLDKLYNPADCSGLHIFALNALRENPNNSNNSALSLKYSASKYNI 255  
 DB 159 KPNSTYSRSDVLYTFANCGSLDILFELNALNTADLQNNSSNAQLLDVCSKGYNI 218  
 QY 256 SWELGNENPNRYTMGRANVSQGLKDYIQLKSLLOPIRIRYASLYGPNIGRPKXVIA 315  
 DB 219 SWELGNENPNRYTMGRANVSQGLKDYIQLKSLLOPIRIRYASLYGPNIGRPKXVIA 315  
 QY 316 LLDGEMKAVGTYDANTVMOHCYIDGRVYKVMDFLTKRLDLTLDSDIRIKIOKVVNTYTPGK 375  
 DB 278 MLKSLFKAGGEIVDSVTMHYVLTNGRTAREDFLNPVDVLDIFISSVQKVPVVESTREK 337  
 QY 376 KWLBSGVVTSAGGTTNLSDSYAAGFLMNTLGMLANOGIDVYIRHSPFDGNYHLVYON 435  
 DB 338 KWLBSGVVTSAGGTTNLSDSYAAGFLMNTLGMLANOGIDVYIRHSPFDGNYHLVYON 435  
 QY 436 FNPDPYMLSLYKRLIGPKVLAHVAGLQKRPGRVIRDLKRLIYAHCTNNHNHNVYRG 495  
 DB 398 FNPDPYMLSLYKRLIGPKVLAHVAGLQKRPGRVIRDLKRLIYAHCTNNHNHNVYRG 495  
 QY 496 STTLFTINHRKRIKLAGTRDKLVHGYLLQPYGOEGLKSKSVQNLQNPVMDGTL 555  
 DB 449 DTLVAINHANTYKRLIPYPPSNKQVDKYLRLPGLGSLSKSVQNLQNLKMDDTL 508

QY 556 PELKPRPLRAGRTLVIPVTMGVYVKNVNALAC 589  
 DB 509 PPLMEKPLRPGSSLGIPAFSYSPFVIRNKAAC 542

RESULT 10  
 ID 09MY10 PRELIMINARY; PRT; 545 AA.  
 AC 09MY10;  
 DT 01-OCT-2000 (T-EMBLrel. 15, Created)  
 DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)  
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)  
 DE Hepatopane.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;  
 OC Bovidae; Bovinae; Bos.  
 NC NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Tissue=Placenta;  
 RA Kizaki K., Nakano H., Takahashi T., Imai K., Hashizume K.;  
 RT "Expression of Hepatopane mRNA in Bovine Placenta During Gestation.";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF281160; AAF87301.2; -  
 DR InterPro; IPR005199; Glyco\_hydro\_79N.  
 DR Pfam; PF03662; Glyco\_hydro\_79n; I.  
 SQ SEQUENCE 545 AA; 61076 MW; FAC4BDFD855B933 CRC64;

Query Match 36.9%; Score 1138; DB 6; Length 545;  
 Best Local Similarity 42.9%; Pred. No. 1, 6e-81;  
 Matches 248; Conservative 85; Mismatches 201; Indels 44; Gaps 10;

QY 20 ACIAPGATLALHLHSLSSQAGRRP-----LPDRAGLKEKTLILDVSTKNPRT 73  
 DB 3 ACRKPG-LRPPILLPLLPGLPDPSPGPAAPADDAE-----LEFFTERPLHL 53  
 QY 74 VNEFNLSGLDPSIIHD-GWLDPLSKRLVTLARGLSPALFPGGKRTDPLQFONLNPARG 132  
 DB 54 VSPAFSLFTDANLADTPRFLILGSPKRLTLARGLSPALFPGGKRTDPLFDPKKEBA 113  
 QY 133 KSRGGPDPYLLKNVEDDIVRSVALDKQCKCIAQHP-DVNLVLOREKAAQMHVLVLEK 191  
 DB 114 FEB-----RSYMLQSQNOI-----CKSGSIPDVEEKLLEWFOEOL-LNREHYQK 156  
 QY 192 QFSNYSNLTILTRSLDKLYNPADCSGLHIFALNALRENPNNSNNSALSLKYSASK 251  
 DB 157 QYCKFTNYSRSDVLYTFANCGSLDILFELNALNTADLQNNSSNAQLLDVCSKGYNI 218  
 QY 252 KYNISWELGNENPNRYTMGRANVSQGLKDYIQLKSLLOPIRIRYASLYGPNIGRPKXVIA 315  
 DB 217 NYNISWELGNENPNRYTMGRANVSQGLKDYIQLKSLLOPIRIRYASLYGPNIGRPKXVIA 315  
 QY 312 NYNISWELGNENPNRYTMGRANVSQGLKDYIQLKSLLOPIRIRYASLYGPNIGRPKXVIA 315  
 DB 276 NYNISWELGNENPNRYTMGRANVSQGLKDYIQLKSLLOPIRIRYASLYGPNIGRPKXVIA 315  
 QY 372 TPQKRLVBSGVVTSAGGTTNLSDSYAAGFLMNTLGMLANOGIDVYIRHSPFDGNYHLVYON 435  
 DB 336 RPLKRWLBSGVVTSAGGTTNLSDSYAAGFLMNTLGMLANOGIDVYIRHSPFDGNYHLVYON 435  
 QY 432 VDNQFNLPYMLSLYKRLIGPKVLAHVAGLQKRPGRVIRDLKRLIYAHCTNNHNHNVYRG 495  
 DB 396 VDNQFNLPYMLSLYKRLIGPKVLAHVAGLQKRPGRVIRDLKRLIYAHCTNNHNHNVYRG 495  
 QY 492 YVGSITLFTINHRKRIKLAGTRDKLVHGYLLQPYGOEGLKSKSVQNLQNPVMDGTL 555  
 DB 447 YKESDITLVAINHANTYKRLIPYPPSNKQVDKYLRLPGLGSLSKSVQNLQNLKMDDTL 508  
 QY 552 DGTLPKPRPLRAGRTLVIPVTMGVYVKNVNALAC 589  
 DB 507 EGTLPALTEKPLRPGSSLGIPAFSYSPFVIRNKAAC 544

```

RESULT 11
ID 090YKS PRELIMINARY; PRT; 523 AA.
AC 090YKS;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Heparanase.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11387326;
RA Goldsmith O., Zeharia E., Aingorn H., Guatta-Rangini Z., Atzmon R.,
RA Michal I., Becker I., Mitrani E., Vlodavsky I.,
RT "Expression Pattern and Secretion of Human and Chicken Heparanase Are
RT Determined by Their Signal Peptide Sequence."
RL J. Biol. Chem. 276:29178-29187(2001).
DR EMBL: AY037007; AA062548.1; -
DR InterPro: IPR005199; Glyco_hydro_79N.
DR Pfam: PF03662; Glyco_hydro_79n; I.
SQ SEQUENCE 523 AA; 58386 MW; 8EB0B7B18C9BF881 CRC64;

Query Match 33.5%; Score 1033.5; DB 13; Length 523;
Best Local Similarity 40.8%; Pred. No. 2.9e-73;
Matches 233; Conservative 82; Mismatches 199; Indels 57; Gaps 11;

QY 27 LVALILHLSLSQAGRRPLPYDRAAGLEKKTLLDVSSTKNPFTVNEFTSLQDPS 86
DB 1 MVLVLLLVLLA-----VPRRTABIQ-----GLRPPICAVSPAFSLTLDAS 44

QY 87 IIND-GLMDFLSKRLVTLARGLSPAFKFGKRTDFLOFQNLNPAKSRGGGPDYLYK 145
DB 45 LAMPDFVALIRPKLHTLASGLSPGLRFGSTPFLF---NPKD-----S 90

QY 146 NYEDIVRSVDALDKQCK-----IAQHPDVLV-----LQREKAAQMLVLLKEQFSNTYS 198
DB 91 TWEKVLSEFOAKDV---CEAWPSFAVVPRLTLTQWPLQEK-----LLLAHSHWKKK 140

QY 199 NILTLARSIDKLYNPDCCGLHLIFALNLRPNNSNSSLSTLXKSASKKNISWE 258
DB 141 NTITSTSLDILHTPSSSGFLVFGNLALRRAGLQWSSNAKOLLGYCAQSYNISWE 200

QY 259 LGNEPNRYRTMGRAVNGSOLGQDYIOLKSLQPIRIYSRASLYGPNIGRPRKNVITALD 318
DB 201 LGNEPNSFRKSGICIDGFLGRDPYHLRQLSQHPLYRHAELVGLDVQGRKHQHLAR 260

QY 319 GPKKVASGTYDAVWQHCYIDGRVVKMDFLTKRLDLTSDQIRKIQKVVNTYTPCKKIY 378
DB 261 SFKSGSKALDSVTMHHYVNGSAREPDLSPEDVDSFATLHDVLGIVEATVPCKKW 320

QY 379 LBEVVTTSAGTNNLSDSYAAGFLMNTLGLMNLNOGIDVIRASFPHGYNHLVDQNFNP 438
DB 321 LGSTGSAVGGGAPQLSTVYAGFMMLDKGLAARIGIDVYMRQVSGAGSYHLVDGFRP 380

QY 439 LPTWYSLLYKRLIGPVYLAHVYAGLQKRPGRGVTRDKRLIYAGCTNHNHANYVAGSIT 498
DB 381 LPTWYSLLYKRLIGPVYLAHVYAGLQKRPGRGVTRDKRLIYAGCTNHNHANYVAGSIT 431

QY 499 LFTINHSRKKITLQAGTADKLVHGYLLQYGOEGLSKSVOLNGLQVWVDGDTLPFL 558
DB 432 LFTINHSRKKITLQAGTADKLVHGYLLQYGOEGLSKSVOLNGLQVWVDGDTLPFL 491

QY 559 KRPDLPAQRTLVLPVTMGFYVKNVNAALAC 589
DB 492 HEMALAPGSTLGLPAPSGFYVIRNAKAIAC 522

RESULT 12

```

```

08T108
ID 08T108 PRELIMINARY; PRT; 515 AA.
AC 08T108;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Heparanase-like protein.
GN BHEPFA.
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
OC Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=p50; TISSUE=posterior silk gland;
RA Koike Y., Shimada T., Suzuki M.G., Mita K., Abe H., Maeda S.,
RA Oseogawa K., deJong P.O.;
RT "Genomic sequence of 320kb containing a kettin orthologue on the Z
RT chromosome in Bombyx mori."
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB079860; BAB85191.1; -
DR EMBL: AB090307; BAC10612.1; -
DR InterPro: IPR005199; Glyco_hydro_79N.
DR Pfam: PF03662; Glyco_hydro_79n; I.
SQ SEQUENCE 515 AA; 59769 MW; FB810ABE6EDDAB CRC64;

Query Match 20.2%; Score 625; DB 5; Length 515;
Best Local Similarity 29.8%; Pred. No. 7.5e-41;
Matches 166; Conservative 90; Mismatches 211; Indels 90; Gaps 17;

QY 60 LILLDVSTKNPVR---TVNEFTSLQDPSLIHDGMLDF-----LSSRLV 102
DB 21 MLVLFVKNGGVRFFVFINEN---QEDIKLISDFSGIDTIEINRINYSDFRLR 76

QY 103 TLARGLSPAFKFGKRTDFLOFQNLNPAKSRGGGPDYLYKNYEDIVRSVDALDKOK 162
DB 77 ELAALSPARLRIGTYSERLIFSKENIPISCH----- 109

QY 163 GCKIAQHPDVLVLOREKAAQMLVLLKEQFSNTYSNLITLARSIDKLYNPDCCGLHLI 222
DB 110 NCSYKSYKSLCOLIEKPKCKHKEFL-----PFIIMGNEMNQINDFCRTNKLTL 160

QY 223 FALNALRPNNNSNSSLSTLXKSASKKNISWEI LGNEPNRYRTMGRAVNGSOLGQ 282
DB 161 FSLNMLNRD-NHGMKRNARLEIFSKKQYALDMQNGNEPNSQHPFNESVTQQLAKD 219

QY 283 YIOLKSLQPIRIYSRASLYGPNIGRP---RKVNIALLDGFMKVASGTYDAVWQHCYID 339
DB 220 FEKLRKILNH-NGYRSHLVIPDTPRQHPRECLAKYIEFLGNGSHYINRSHWQYLLN 278

QY 340 GRVYKVMDELTKRLDLTSDQIRKIQKVVNTYTPCKKI--WLEBVVTTSAGTNNLSDSY 397
DB 279 STYALGLEDNFKNPETFDLKRQIETWQNTKTY---KAIPLMLSTSSSYGGAAGLSNTY 335

QY 398 AAGFLMNTLGLMNLNOGIDVIRASFPHGYNHLVDQNFNPLDPYMLSLYKRLIGKVL 457
DB 336 AAGFLMIDKLGSAKNYSITVIRQSFIG-GYSLVDENLKPLPDMWISVLYKRLVGNKVL 394

QY 458 AY--HVAQLOKRPGRGVTRDKLIYAGCTNHNHANYVAGSITLFTINHSRKKIKLAG 515
DB 395 QVQNCNSRFQ-----RLYTHCTNRKTYNT-LSAVLYGVNLEKARFPFLNG 440

QY 516 TL---RDKLVHGYLLQYGOEGLSKSVOLNGLQVWVDGDTLPFLPAQRTLVIP 572

```



Db 441 TAIHGDLLIHEYLIA-DSNNKSKTILLNGMPLY--ESMLHINRPHRYGRVSLP 497  
 Qy 573 PTMGFYVYVNNALAC 589  
 Db 498 PYSIGFWIKKTSITVC 514

## RESULT 13

Q8H615 PRELIMINARY; PRT; 544 AA.

AC Q8H615;  
 DT 01-MAR-2003 (TEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)  
 DE OSJNB0035103.11 protein.  
 GN OSJNB0035103.11  
 OS Oryza sativa (Japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OC NCBI\_TaxID=3947;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GAS) genomic DNA, chromosome 6, BAC  
 clone:OSJNB0035103."  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF003019; BAC2238.1;  
 DR Interpro; IPR005199; Glyco\_hydro\_79N.  
 DR Pfam; PF03662; Glyco\_hydro\_79n; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 SQ SEQUENCE 544 AA; 58723 MW; ECDE695F0E22A269 CRC64;

Query Match 12.4%; Score 384; DB 10; Length 544;  
 Best Local Similarity 24.8%; Pred. No. 1.3e-21;  
 Matches 163; Conservative 81; Mismatches 198; Indels 216; Gaps 26;

Qy 25 GALYLAALLHLSTSQAGDRRLPVRAAGLKEKTLILDVSTKPFRTVNEFLSLQID 84  
 Db 9 GLCFILHL-----RLGGAFAVGVAGVDRRAIAATDEDFVCAAMD 50  
 Qy 85 ---PSIHDMDLF-----LSKRLVTLARGLSPATL 113  
 Db 51 WPPDKCDYTCGWSGLASLNLVENSFLQISKFSPPPPPDLSNKLINAIKAFPLKL 110  
 Qy 114 RFGSKRTDFQONTLRNPAKSGAP--GPDYLLKNVEDDIYRSDVALDKQKCKIAQHPDV 172  
 Db 111 RLGGSLQDQKLVY-----GTGDSGGPCAP--FVKN----- 137  
 Qy 173 MLVLOREKAQWELVLKEQFSNTYSNLITARSIDLKYNFADCGSLHLIFALNALRN- 231  
 Db 138 -----TSEWFGFTQGLPL--HRMDLNAFFQKSGARIVFGINALNGRV 179  
 Qy 232 --FNNS---WSSSALSLIKYSASKKYN--SWEIENEPNNYRTMGRAVNSQLGKDYI 284  
 Db 180 PLVDGSGWGWDTYNAASLRTYASKGYKIHGHELNELSGSGV--GTVGADQYADVI 237  
 Qy 285 QLSLQPIRIYRASLYGPNIGRPRKNVIALDGMKVASSTVDAVTWQHCYIDGRVYK 344  
 Db 238 ALKSLVD--TIYO-----GNPSKPLVLAQGFED--AGMFEVI----- 272  
 Qy 345 VMDFLKTR--LIDPLISDQIRKIQVNTYTPGKKI--MLEGVYT-----SAG-- 388  
 Db 273 ---VKTRPMLNVNTHITNLGPGVDTHILEKILNPSTYDGNVSTFNSLQGLISAGTS 328  
 Qy 389 -----GTNNLSDSYAAGFLMNTLGLANOGIDIVYIRHSFPHGYNHLVDQ 434  
 Db 329 AVAMVGSAGAYNSGRHLVTDSEVFSFWFLDQGMASKYDTKSYCGSLIGGNYGLNKE 388  
 Qy 435 NENPLDPYMLSLYKRLIGPKYLAHVAGIQKRPGRVIRDLRLYACVCTHHNNHNYR 494  
 Db 389 TFDENPDIYSLGMLNHRMLNGTGLSATPNG-----TNMRTYAHQAMD-----S 431

Qy 495 GSTILFIINL-----HRSKKIKYLAAGTLDDKLVHGYLQIPYGO 532  
 Db 432 PGITLLILINSGNTSQSVTSSEGANHYKKSRTKRLHLAGSMBE---EYHLT--AK 485  
 Qy 533 EG-LKSKVQNGOPLVWVDGTLPELKPRPLRAGRTLVIPPTMGFYVYVNNALAC 589  
 Db 486 DGSLSQVWMLNGRALVADENGETPRLEPKVDAQPIAVAPISYIFAHINHPAPAC 543

## RESULT 14

Q9SDA1 PRELIMINARY; PRT; 521 AA.

AC Q9SDA1;  
 DT 01-MAY-2000 (TEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN F13G24.30.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_TaxID=3702;  
 RN  
 RP SEQUENCE FROM N.A.  
 RA Beyan M., Van der Schueren J., Chuang Y.J., Voet M., Robben J.,  
 RA Volckaert G., Bancroft I., Mewes H.W., Lemcke K., Meyer K.F.X.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RN  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL133421; CAB62595.1;  
 DR PIR; T45608; T45608.  
 DR GO; GO:0004285; P:trypsin activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR Interpro; IPR005199; Glyco\_hydro\_79N.  
 DR Interpro; IPR001254; Peptidase\_S1.  
 DR Pfam; PF03662; Glyco\_hydro\_79n; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Hypothetical protein

Query Match 12.4%; Score 382; DB 10; Length 521;  
 Best Local Similarity 25.8%; Pred. No. 1.5e-21;  
 Matches 154; Conservative 76; Mismatches 187; Indels 180; Gaps 28;

Qy 75 NENFLSLQID--PSIHD-----GWLDFLS--SKRLVTLA--RGLSPAFIRFGSKR 119  
 Db 22 DENFVCATLDWP--HDKCNVDCCPWGYSVIMDLRPLRLTKALAFKPLRIRIGSL 78  
 Qy 120 TDFLQFQ--NLNPA-----KSRGGPDPYLLKNVEDDIYRSDVALDKQKCKIAQHPDV 173  
 Db 79 QDQYIVDGNLKTCPRFQKNSG-----LFGSKGCG----- 110  
 Qy 174 LVLOREKAQWELVLKEQFSNTYSNLITARSIDLKYNFADCGSLHLIFALNALR- 229  
 Db 111 -----LHM-----KRWDELNLSLTKAGAVVYGLNALRNRHK 142  
 Qy 230 --FNPNNSWSSSALSLIKYSASKKYN--SWEIENEPNNYRTMGRAVNSQLGKDYI 285  
 Db 143 LRGAAGAMOHINTOPFLNVTYASKGVVIDSWERGNELSGSGV--GASVSAELYGKDLIV 200  
 Qy 286 LKSLQPIRIYRASLYGPNIGRPRKNVIALDGF-----MKAG--STYDAVTWQH 335  
 Db 201 LKDVIN--KYKSNWLEKPLIVAP-----GGFYEQQYTKLSTSGSVVDVVT--HH 249  
 Qy 336 CYIDGR-----VYKWDPLKTRLLDPLISDQIRKIQVNTYTPGKKI--MLEGVYT--SAG- 388  
 Db 250 IYNGSGNDPALVYKIND---PSYLSQVSKTFPVQVNTIGEHG---WASPMVGSAG 302  
 Qy 389 ---GTNNLSDSYAAGFLMNTLGLANOGIDIVYIRHSFPHGYNHLVDQFNPPLPYWS 445



Db 303 YNSGRHVSDFIDFWLQDQMSABHNTKVYCRQTLVGGFYGLEKGTFFVNPDIYSA 362  
QY 446 LLYRRLIGPKYLAHVAGLQKRPGRVIRDKLRIYAHCTNNHNNVYRSGITLFINLH 505  
Db 363 LLMHRLMGKVLAVQTDG-----PQQLVYAHCSKG-----RAGVTLLILNLS 405  
QY 506 -----RSRKKIKLAGTLRDLV-----HQYLLQPYGOE 533  
Db 406 NOSDFTVSVNGINVLNABSRKKSLDITLKRPFMSWIGSKASDGYLNBEETHLP--EN 463  
QY 534 G-LKSKSVQNLGQPLVMVDGTLPELKRPLRAGRTLVIPVTMGFYVYKAVNALAC 589  
Db 464 GVLRSKTMVLNKGSKLKPATGDIPLSEPLVRSVNSPLNVLPLMSFTVLPRPDAAC 520

RESULT 15  
Q9FF10 PRELIMINARY: PRT: 543 AA.  
AC Q9FF10: 01-MAR-2001 (TRENBLrel. 16, Created)  
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Similarity to heparanase.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Columbia;  
RX MEDLINE=97471969; PubMed=9330910;  
RA Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,  
RA Miyajima N., Tabata S.;  
RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence  
RT features of the 1.6 Mb regions covered by twenty physically assigned  
RT clones.";  
RL DNA Res. 4:215-230(1997).  
DR EMBL; AB005249; BAB09947.1; -;  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR005199; Glyco\_hydro\_79N.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR Pfam; PF03662; Glyco\_hydro\_79n; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
SO SEQUENCE 543 AA; 60250 MM; 0FA2248948282FF6 CRC64;

Query Match 12.4%; Score 382; DB 10; Length 543;  
Best Local Similarity 25.8%; Pred. No. 1.6e-21;  
Matches 154; Conservative 76; Mismatches 187; Indels 180; Gaps 28;

QY 75 NENPLSLQD--PSLIH-----GMDPLS---SKLVTLA-RGLSPALRFGGR 119  
Db 44 DENFVCAITLDMWP---HDKCNVDOCPWGSVIVNNILTRPLTKAIRKPLRIRIGSL 100  
QY 120 TDFLQFC--NLRNA---KSRGGPDPYLLKNYEDIVRSVDVALDKQGGCKIAHPDM 173  
Db 101 QDOYIVDVGNLKTCPRPFGKXNSG-----LFGSKG----- 132  
QY 174 LVLQREKAAQWHLVILKEQFSNTYSNLILTRASLDKLYNPADCGSLHIFALNLR----- 229  
Db 133 -----LHM-----KRWDELNSFLTATGAVVTFGLNALRHRK 164  
QY 230 ---RNPNSNNSSSLSLTKYSASKYNI-SWEELNEPNRYTWGRAVNGSOLGKDIQ 285  
Db 165 LRGRAMGAMDHINTQDPLANTVSKGYIDSMFENELSSGV--GASVSAELYGKDLIV 222  
QY 286 LKSLIQPIRIYSRSLYGNITGRPRKNVIALDGF-----MKVAG-STDAVTWQH 335  
Db 223 LKDVIN--KVKYNSMLHKKPIIIVP-----GGFYEQQWYTKLIRISGSPVDVVT--HH 271  
QY 336 CYIDR-----VYVMDPELKTRLDITLSDQIRKIQKVVNTYTPGKKIWLGVVTTTSAG- 386

Db 272 IYNLGSNDPALVKIMD---PSYLSQVSKTEKVDVNOITIOEHP-----WASPMWGESGGA 324  
QY 389 ---GTNNLSDSYAAGFIMNLTGLMANGIDIVYIRHSFFDGNVNLVDQNFPLJEDYLS 445  
Db 325 YNSGRHVSDFIDFWLQDQMSABHNTKVYCRQTLVGGFYGLEKGTFFVNPDIYSA 384  
QY 446 LLYRRLIGPKYLAHVAGLQKRPGRVIRDKLRIYAHCTNNHNNVYRSGITLFINLH 505  
Db 385 LLMHRLMGKVLAVQTDG-----PQQLVYAHCSKG-----RAGVTLLILNLS 427  
QY 506 -----RSRKKIKLAGTLRDLV-----HQYLLQPYGOE 533  
Db 428 NOSDFTVSVNGINVLNABSRKKSLDITLKRPFMSWIGSKASDGYLNBEETHLP--EN 485  
QY 534 G-LKSKSVQNLGQPLVMVDGTLPELKRPLRAGRTLVIPVTMGFYVYKAVNALAC 589  
Db 486 GVLRSKTMVLNKGSKLKPATGDIPLSEPLVRSVNSPLNVLPLMSFTVLPRPDAAC 542

Search completed: May 6, 2004, 13:45:40  
Job time: 47 secs



```

Db      708 GGGAGCTCTGAGCCAGACACCCCTGTACCGACAGCTGAGCTGTACGGCTTGACCTGG 767
Qy      920 GGGGGCCGAGAGAAATGATGATGCTCCCTCTTATGATGATGATGAAAGTGGCGAGAAATA 979
Db      768 GGGAGCCCCCGACAGACACCCAGACCTGCTCAGAAAGCTTCAAGAAATCTGAGGGAGAA 827
Qy      980 CAGTATGAGATGAGTACCTGCGCAATGCTACATGATGAGCGGGGTGCTCAAGGTATGG 1039
Db      828 CGATTGCTGTGCTGACCTGGACCACTACTATGTAATGAGCCAGAGTGCACAGAGAG 887
Qy      1040 ACTTCTGAAAATCTGCTGTGTAACAACACTCTGACACGATTTAGAAAATTCAGAAA 1099
Db      888 ATTTCCTGAGACCTGTAAGTCTGAGACTCTTTGCTCACTGCTACACAGATGCTCTGGGA 947
Qy      1100 TGGTAAATACATACCTCCAGAGAAAGAAATTTGGCTTGAAGTGTGTGACCACTCG 1159
Db      948 TCGTGGAGCAACGGTCCCGGCAAGAGATGTGGTGGTGAACCGGCTCGGCTTACG 1007
Qy      1160 CTGAGGCAACAAATCTATCCGATTCCTATGCTGACGAGATCTTATGTTGAAACATT 1219
Db      1008 GCGGGGGGGCCCCAGCTCTCAACACTATGTGGCGGCTTATGTGGCTGGAACAG 1067
Qy      1220 TAGGAATGCTGGCAATCAGGCGATGATGTGTGATAGGCACTCATTTTGAACATG 1279
Db      1068 TGGGGTTGGCGCTCGGCGTGTGCAATGATGTGTGATGAGGAGGTCTCTTTGGTGTG 1127
Qy      1280 GATACATCACTCTGTGACAGAAATTTTAACCATTTACGACACTGCTGCTCTCTCC 1339
Db      1128 GAGCTATCACTGTGTGATGTCGGGCTTCAAGGCTTTGCGGACTCATGCTGTCTCTCC 1187
Qy      1340 TCTAACAGGCTGTATGCGCCCAAGTCTTGGCTGTGATGTGCTGGGCTTCCAGGGA 1399
Db      1188 TATACAAAGAGCTGTGTGGGCAACCGGTACTACAGGCAAGCTGTGAGCAAGCGATGCC 1247
Qy      1400 AGCCAGCGCTGCGCGAGTATCCGGGCAAACTAAGATTTATGCTCACTGACAAAC 1459
Db      1248 GGGGCGCGC-----GGGTCTACCTGCACTGACCAACC 1280
Qy      1460 ACCAACAACAACCTACGTTCTGTGGTCCATTAACCTTTTATCACTCACTTGATCAT 1519
Db      1281 CCGGCAACCCCAATACCGGGAAGGAGATGTGACACTGTTGCTTGAACCTCTCCAG 1340
Qy      1520 CAGAAAGAAATCACTGCTGCTGGGACTCTCAGAGCAAGCGTTCACGATACCTG 1579
Db      1341 TGACCCAGAGCTGTGACCTCTTACAGAGTGTGTGAGTAAAGTGTGATCACTTCTCC 1400
Qy      1580 TGAAGCTTATGGGCAAGAGGCTTAAAGTCAAGTCACTGCACTGAAATGGCCAGCCCT 1639
Db      1401 TGTGCCCCCAGGCAAGAGACATCTGTCCAGAGAGGTGACGTGAATGGCCGCTAC 1460
Qy      1640 TACTGATGTGTGAGAGCGGAGCTCCAGAAATTTGAAGCCCGGCTTGGGCGCGGC 1699
Db      1461 TGCAGATGTGTGAGAGAGAGACCTCCCGGCTGCAAGATGAGCCCTTGGCCCTG 1520
Qy      1700 GACATTTGATCACTCTCCAGTCAACATGAGCTTTTATGTGTGATGATGATGCTT 1759
Db      1521 GCAAGCTCGGCTGCGAGCTTCTTACAGTGTTCACAGATCAGAAAGCTTAAAGCTTA 1580
Qy      1760 TGGCTGCG 1767
Db      1581 TTGCTTGC 1588

```

## RESULT 2

```

US-09-435-739-43
; Sequence 43, Application US/09435739
; Patent No. 6664105

```

## GENERAL INFORMATION:

```

; APPLICANT: Pecker, Itis
; APPLICANT: Viodevsky, Israel
; APPLICANT: Feinstein, Elena
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY
; TITLE OF INVENTION: EXPRESSION OF SAME IN GENETICALLY MODIFIED CELLS

```

```

; FILE REFERENCE: 00/20454
; CURRENT APPLICATION NUMBER: US/09/435,739
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 43
; LENGTH: 2396
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-435-739-43

```

```

Query Match      15.9%; Score 282.6; DB 4; Length 2396;
Best Local Similarity 52.8%; Pred. No. 3e-77;
Matches 677; Conservative 0; Mismatches 574; Indels 30; Gaps 2;

```

```

Qy      474 AGATTAACAGAAAGCTGCAAGATTGCGCAGACACCTGATGTTATGCTGTGCTCCAAAG 533
Db      932 AGTCAACATGATATTTGCAAGTCTGAGCCGCTCTGCTGCGGTGTGAGAAACTCCA 991
Qy      534 GGAAGAGCAGCTCAGATGATCTGTTCTTCTTAAAGAGCAATTCATACTTACAG 593
Db      992 GGTGAATGSCCTTCCAGAGAGCTGTGTGCTCTCCAGAGCAGTACCAAAAGAGTTCA 1051
Qy      594 TAATCTCATTAATTAACAGCAGCTCTTCAAGCAAACTTATTACTTGTGTGCTGTG 653
Db      1052 GAAACGACCTTACTCAAGAAAGCTCAGTGAATGCTCTACAGTTTGGCCAAAGTCTCGGG 1111
Qy      654 ACTCACCTGATATTTGCTCTTAAATGACAGCGCGTGTATCCCAATTAATCTCGAAG 713
Db      1112 GTTACGCTGATCTTTGTGTCTTAAATGCGTTACTACAGAACCCAGACTTACGGTGA 1171
Qy      714 TTCTAGTCCCTGAGTGTGTTGAAGTACAGCGCCAGCAAAAGTACAACTTTCTGGGA 773
Db      1172 CTCACAGCCAGCTTCTCTTCTGACTACTGCTCTTCCAGAGTTATTAATCTCTGGGA 1231
Qy      774 ACTGGTATGAGCAATAACTATCTATGAGACATGACATGCGCGGCAAGTAATGAGAGCA 833
Db      1232 ACTGGCATATGAGCCACAGCTTCTGAGAAAGCTCAATCTCTCATGATGGGTTGCA 1291
Qy      834 GTTGGAAAGGATTACATCCAGCTGAAGAGCTGTGTGCAAGCCATCCGATTTATTCAG 893
Db      1292 GTTAGGAAAGACTTTGTGGTGTGCAATACTTCTCA---AAGTCACTTTCACAAA 1348
Qy      894 AGCCAGCTTATATGCGCTTAATTTGGCGCGGCGAGAGAAATGTACTGCGCCTCCAGA 953
Db      1349 TGCAAAACTATATGCTGTGACATCGGTCAAGCTTCCAGGAGACAGTAACTGCTGAG 1408
Qy      954 TGAATTCATGAGGTGCGAGAGAGTACAGTATGACATTAAGTCACTGCACTTGTCAAT 1013
Db      1409 GAGTTTCTGTAAGGCTGCGGAGAGAGATGATCGACTGTCTTACATGGAATCACTATTACT 1468
Qy      1014 TGAATGCGCGGTGTGATCAAGTATGATGATCTCTGAAAATCTGCGCTGTATACACACTCTC 1073
Db      1469 GAATGAGAGATGCTCTCAAAAGAAATTTTGTAGTGTGATGCGGTGAACCTTTAT 1528
Qy      1074 TGACCAATTAAGAAATTCAGAAAGTGTAAATACATCACTCCAGAAAGAAAGATTGG 1133
Db      1529 TCTCTGTGCAAAAATTTCTGAAGGTCACTAAAGATCAACCTGTGCAAGAAAGCTG 1588
Qy      1134 GCTTGAAGGTGTGTGACCACTGAGTGAAGGCAAAACAAATCTATCCGATTCCTATG 1193
Db      1589 GTTGGAGAGAGAGCTCAGCTTACGTTACGGTGGCGGTGACCTTGTGCAACCTTTG 1648
Qy      1194 TGACAGATTTCTTATGTTGAACCTTTAAGATGCTGGCAATCAGGGCAATTAATGCTG 1253
Db      1649 AGCTGCTTATATGCTGTGATTAATTTGGGCTGTGACCCAGTATGAGTAAAGTGTG 1708
Qy      1254 GATACGCGCATCATTTTGTGACATGATGATCAATCACTGTGTGACCAAAATTTTAAACC 1313
Db      1709 GATAGAGAGGTGTCTTCCAGAGAGCAACTACATTAATGATGATGATGATGATGAGCC 1768
Qy      1314 ATTACCAAGTACTGCTCTCTCTCTCAAGCGGCTATGCGGCCCAAGCTCTTGC 1373

```

Db	1769	TTTACCTGATTAACAGGCTCTCTCTTCTGTTCAGAAAACCTGGTAAGTCCAGGGTGTACT	1828
QY	1374	TGTGCATGTGCGCTGGGCTTCACGGGAAGCACGGCGCTGGCCGATGATCCGGACAACT	1433
Db	1829	GTCAAGAGTTGAAAGGCCAGA-----CAGAGCAAACT	1861
QY	1434	AAGATTTATAGCTCATCTGCACAAACCAACCAACCAACTACGTTGGTGGTCCATPAC	1493
Db	1862	CCGAGTGTATCTCCACTGCACTTAACGTTATATACCAACGATATCAGAAAGAGATCTAAC	1921
QY	1494	ACTTTTATATCACTCACTTGCACTGCATCAAGAAAGAAAATCAAGCTGGCTGGGACTTCAG	1553
Db	1922	TCTGTATGTCTGAAACCTCATTAATGTACACAAAGCACTGAAGGTAACGGCTCCGTTGTT	1981
QY	1554	AGACAGCTGTTTCCACCACTACCTGTGACGCCCTATGAGGCGAGGGGCTAAAGTCCA	1613
Db	1982	CAGGAAACGAGTGGATAGTATCTTCTGAAGCTTCGGGGCCGAGTGAATTACTTTTCCA	2041
QY	1614	GTCAGTGCACATGAATGGCGACGCCCTTAGTGTATGTGACGACGAGACCCCTCCAGAAAT	1673
Db	2042	ATCTGTCCAATCGAAGGGTCAATTTCTGAAGATGGTGAATGACGACCCCTGCCAGCTTT	2101
QY	1674	GAAAGCCCGCCCCCTTCGGGGCCGCGACATTGGTCACTCCTCCAATCACCATGGGCTT	1733
Db	2102	GACAGAAAAACCTTCGCCCGCAGGAAGTCACTAAAGCTGCCTTTTCTTATGATTT	2161
QY	1734	TTATGTGTCAGAAATGTCAA	1754
Db	2162	TTTTGTCAATAGAAATGTCCA	2182

```

RESULT 3
US-09-435-739-45
: Sequence 45, Application US/09435739
: Patent No. 6664105
:
: GENERAL INFORMATION:
: APPLICANT: Pecker, Irit
: APPLICANT: Vladavsky, Israel
: APPLICANT: Feinstein, Elena
: TITLE OF INVENTION: POLYPEPTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY
: TITLE OF INVENTION: EXPRESSION OF SAME IN GENETICALLY MODIFIED CELLS
: FILE REFERENCE: 00/20454
: CURRENT APPLICATION NUMBER: US/09/435,739
: CURRENT FILING DATE: 2001-06-05
: NUMBER OF SEQ ID NOS: 47
:
: SOFTWARE: Patentin version 3.0
:
: SEQ ID NO 45
:
: LENGTH: 2396
:
: TYPE: DNA
:
: ORGANISM: Mus musculus
:
: FEATURE:
:
: NAME/KEY: CDS
:
: LOCATION: (594)..(2198)
:
: US-09-435-739-45

```

	Query Match	Similarity	Score	DB	length
Best Local	15.9%	52.8%	287.6	4	2396
Matches	677	Conservative	0	Mismatches	574
				Indels	30
				Gaps	2
QY	474	AGATTAACAGAAAGGCTGCAAGATTGCCAGCACCCTGATGTTATCTGTGCTCCAAAG	533		
Db	932	AGTCACCACTGATATTGTCAGGCTGAGGCGGCTCTGCTGCGGTTTGAGGAAATCCCA	991		
QY	534	GGAGAAGGACGCTCAGATCATCTGCTCTTCTTAAAGAGCAATTCCTCAATCTTACAG	593		
Db	992	GGTGGATGGCCCTTCCAGAGCTGTGCTGCTCCAGAGCAATTAACAAAGAGTTCAA	1051		
QY	554	TAACTCATATTAAACGACGAGCTCTAACAACACTTTAACTTGCTGATGCTGTGG	653		
Db	1052	GAAACGACCTACTCAAGAGCTCACTGACACGCTCTACAGTTTGCCAAAGCTCGGG	1111		
QY	654	ACTCACTGATATTGCTCTTAAATGCACTGGCTGTTAATCCCAATTACTCTGGAAACG	713		

[illegible]

RESULT 4  
US-08-922-170B-9  
Sequence 9, Application US/08922170B  
Patent No. 5968822  
GENERAL INFORMATION:  
APPLICANT: Iris Pecker, Israel Vlodavsky and Elena  
APPLICANT: Feinstein  
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE  
TITLE OF INVENTION: HAVING HEPARASE ACTIVITY AND EXPRESSION OF  
TITLE OF INVENTION: SAME IN TRANSFUSED CELLS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Mark M. Friedman c/o Robert Sheinbein  
STREET: 2940 Birchtree lane  
CITY: Silver Spring  
STATE: Maryland  
COUNTRY: United States of America  
ZIP: 20906  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
COMPUTER: Twinhead Slimote-890TX  
OPERATING SYSTEM: MS DOS version 6.2,  
OPERATING SYSTEM: Windows version 3.11  
SOFTWARE: Word for Windows version 2.0 converted to  
SOFTWARE: an ASCII file  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/922,170B  
FILING DATE: 2 SEP 1997  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Friedman, Mark M.  
REGISTRATION NUMBER: 33,863  
REFERENCE/DOCKET NUMBER: 910/1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 972-3-5625553  
TELEFAX: 972-3-5625554  
TELEX:  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1721  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-922-170B-9  
Query Match 15.7%; Score 278.8; DB 2; Length 1721;  
Best Local Similarity 52.4%; Pred. No. 3.8e-76;  
Matches 684; Conservative 0; Mismatches 592; Indels 30; Gaps 2;  
QY 474 AGATAACACAAAGGCTGCAAGATGCGCCAGCAGCCTTGATGTATGCTGTGCTCAAG 533  
DB 425 AGTCAACAGAGATATTTGCAATATGATCCATCCCTCGATGTGAGAGAGAGTTAG 484  
QY 534 GGAAGAGGAGCTCAGATGATCTGTTCTTTAAAGAGCAATTCCTCAATCTTACAG 593  
DB 485 GTTGAGATGCGCTCAACAGAGCAATTCCTCGAGAACCTACAGAGAGAGTTCA 544  
QY 594 TAATTCATATTAACAGCCAGGCTCTGACAACTTTATTAATTTGCTGATGCTCTG 653  
DB 545 GAACAGACCTACTCAGAGAGCTCTGATGATGCTATACACTTTTGCAAACTGCTCAG 604  
QY 654 ACTCACTGATATTTGCTTAATGCACTGCTGATATCCAAATCACTCTGAGAGAG 713  
DB 605 ACTGAGCTGATCTTGGCTTAATGCTTATTAAGACAGCAATTTGAGAGAGAG 664  
QY 714 TTCTAGTCCCTGATGTTGTTGAAGTACAGCCGAGAGAGAGAGAGAGAGAGAGAG 773  
DB 665 TTCTAATGCTCAAGTGTGCTCTGAGCTACTGCTCTTCCAGAGGAGATTAACATTTCTGGGA 724

RESULT 5  
US-08-922-170B-11  
Sequence 11, Application US/08922170B  
QY 774 ACTGGTATATGAGCAAAATATATGACACATGATGCGCGGAGAGTAATGACAGCA 833  
DB 725 ACTAGGCAATGAGACCTTACAGTTTCTTTAAGAGGCTGATATTTTATCAATGAGTGGCA 784  
QY 834 GTTGGAGAGAGATTAATCCAGCTGAAAGCCTGTTGACGCCATCCGATTTTATCCAG 893  
DB 785 GTTAGAGAGAGATTAATTAATGATTAAC---TTCTAAGAGAGATCCACTTCAAAA 841  
QY 894 AGCCAGCTTATGAGCCCTTAATATGAGCGCGGAGAGAGAGATGATGAGCCCTCTAG 953  
DB 842 TGCAAACTCTATGCTCTATGTTGCTGAGCTCGAAGAGAGAGAGAGAGAGAGAGAG 901  
QY 954 TGATTCATGAG 1013  
DB 902 GAGCTTCCTGAG 961  
QY 1014 TGATGCGCGGAG 1073  
DB 962 GAATGAG 1021  
QY 1074 TGACCAATTAAG 1133  
DB 1022 TTCACTGTGCAAAAAGTTTCCAGAGGTTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1081  
QY 1134 GCTTGAAG 1193  
DB 1082 GTTAG 1141  
QY 1194 TGACAGATTTCTTATGTTGAACCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1253  
DB 1142 AGCTGCTTATATGAG 1201  
QY 1254 GATACGAG 1313  
DB 1202 GATAG 1261  
QY 1314 ATTAACAG 1373  
DB 1262 TTACAG 1321  
QY 1374 TGACAG 1433  
DB 1322 GAG 1354  
QY 1434 AAGAGATTAAG 1493  
DB 1355 TGAGATTAAG 1414  
QY 1494 ACTTTTATATCACTGAG 1553  
DB 1415 TCTGATGAG 1474  
QY 1554 AG 1613  
DB 1475 TAAAG 1534  
QY 1614 GTCAAG 1673  
DB 1535 ATCTGCACTCAATGAG 1594  
QY 1674 GAG 1733  
DB 1595 AATGAG 1654  
QY 1734 TTATGAG 1779  
DB 1655 TTGATGAG 1700

Patent No. 5968822  
 GENERAL INFORMATION:  
 APPLICANT: Iris Pecker, Israel Vlodavsky and Elena  
 APPLICANT: Feinstein  
 TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE  
 TITLE OF INVENTION: HAVING HEPARANASE ACTIVITY AND EXPRESSION OF  
 TITLE OF INVENTION: SAME IN TRANSFUSED CELLS  
 NUMBER OF SEQUENCES: 12  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Mark M. Friedman c/o Robert Sheinbein  
 STREET: 2940 Birchtree lane  
 CITY: Silver Spring  
 STATE: Maryland  
 COUNTRY: United States of America  
 ZIP: 20906  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
 COMPUTER: Twinhead\* Slimmode-890TX  
 OPERATING SYSTEM: MS DOS version 6.2,  
 OPERATING SYSTEM: Windows version 3.11  
 SOFTWARE: Word for Windows version 2.0 converted to  
 SOFTWARE: an ASCII file  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/922,170B  
 FILING DATE: 2 SEP 1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Friedman, Mark M.  
 REGISTRATION NUMBER: 33,883  
 REFERENCE/DOCKET NUMBER: 910/1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 972-3-562553  
 TELEFAX: 972-3-562554  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 11:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1721  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 US-08-922-170B-11

Query Match 15.7%; Score 278.8; DB 2; Length 1721;  
 Best Local Similarity 52.4%; Pred. No. 3.8e-76;  
 Matches 684; Conservative 0; Mismatches 592; Indels 30; Gaps 2;

QY 474 AGATTAACAGAAAGGCTGCAAGATTGGCCAGCAACCTGATGTTATGCTGCTCAAG 533  
 DB 425 AGTCAACCAAGATTAATTTGCAAAATATGATCCATCCCTCGATGAGAGAAAGTTAG 484  
 QY 534 GAGAAAGGAGCTCAGATGCACTGCTTTCTTAAAGAGCAATTTCCATATCTTACAG 593  
 DB 485 GTTGAATGAGCCCTACCAAGAGCAATTTGCTTACCCAGAACACTACAGAAAGTTCAA 544  
 QY 594 TAATCTCATATTAACAGCCAGCTCTCTAGCAAACTTATTAACCTTTGCTGATGCTG 653  
 DB 545 GAACAGACCTACTCAAGAAAGCTCTGATGATGCTTATTAACCTTTGCAAACTGCTCAAG 604  
 QY 654 ACTCCAGCTGATATTGCTCTTAATGCACTGCGCTGTAATCCCAATTAACCTCTGAACAG 713  
 DB 605 ACTGCACTTATCTTGGCTTAATATGCTTATTAAGAACAGCAAGATTGGCAGTGAACAG 664  
 QY 714 TTCTAGTCCCTGATGCTGTTGAAGTACAGCGCCAGAGAAATTAACAATTTCTTGGGA 773  
 DB 665 TTCTATGCTCAAGTGTCTCTGACATGCTCTTCCAAAGGATTAACAATTTCTTGGGA 724  
 QY 774 ACTGGGTAATGAGCCAAATTAATCTATGAGCAATGATGAGCCGAGTAATATGAGCA 833  
 DB 725 ACTAGGCAATGAACCTTAACAGTTTCTTAAAGAGGCTGATATTTTATCAATGGGTGCA 784

QY 834 GTTGGAAAGATTACATCCAGCTGAAGAGCTGTGGAGCCCATCCGATTTATTCAG 893  
 DB 785 GTTAGAAGAAATTAATTAATGATGATTAAC---TTCTAAGAAAGTCCACTTCAAAA 841  
 QY 894 AGCCAGCTTAATATGCTCTTAATTTGGCGGCGGAGAGAAATATCATGCGCTCTAGA 953  
 DB 842 TGCAAACTCTATGCTCTGATGTTGGTCAAGCTTCAAGAAAGAGCGCTAAGATGCTGA 901  
 QY 954 TGGATTCAATGAAGGTCAGAGAAATGACATGATGAGTCTGCAATCTGCAATCTGCTCAT 1013  
 DB 902 GAGCTTCTGTAAGGCTGTGTGAGAGATGATGATTAATCACTTACATGATCTACTATT 961  
 QY 1014 TGATGCGGCGGTGTCAAGTGTGATGACTTCTGAATAATCTGCTGTGACACACTCTC 1073  
 DB 962 GAATGAGCGACTGCTACCAAGGAAAGATTTCTAAACCTGATGATATGACATTTTAT 1021  
 QY 1074 TGACCAATTTGAGAAATTCAGAAAGTGTAAATACATCACTCCAGAAAGAAATTTG 1133  
 DB 1022 TTCACTGTGCAAAAAGTTTCCAGTGTGTGAGACACAGGCTGTGCAAGAAAGTCTG 1081  
 QY 1134 GCTTGAAGTGTGTGACCACTCAGCTGAGGACCAAACTATCTATCCGATTTCTATGC 1193  
 DB 1082 GTTAGAAGAAACAAGCTGTGATATGAGAGCGGAGCGGCTGTATCCGACACTTTC 1141  
 QY 1194 TGCAGATTTCTTATGTTGAACACTTGAAGATGCTGCGCATCAGGGCATGATGTCGT 1253  
 DB 1142 AGCTGGCTTTATGTTGCTGATTAATTTGGGCTCTCAGCTTCAAGTGAATGAAAGTGT 1201  
 QY 1254 GATACGCACTCATTTTGTGACATGATACATCACTCTGTGACCAAGATTTTAAACC 1313  
 DB 1202 GATGAGCAAGATTTCTTTGAGAGCAAGAACTACATTTATGATGATGAATAAACTTGCATCC 1261  
 QY 1314 ATTAACAGCTACTGCTCTCTCTCTCTCAAGAGCTGATCGGCCCAAGTCTTGCC 1373  
 DB 1262 TTATCTGATTAATGCTGCTATCTCTCTCTCAAGAAATGCTGTGCGACCAAGGTTAT 1321  
 QY 1374 TGTGATGTTGGCTGGCTCCAGCGAAGCCAGGCTGTGAGTATCCGGACAAACT 1433  
 DB 1322 GCGAAGCTGCAAGTTTCAAGAAAGG-----AGCT 1354  
 QY 1434 AAGATTTATGCTCACTGCACAAAACCAACACCACTACCTGTGTGCTCATTTAC 1493  
 DB 1355 TCGAGTATACCTTCACTGCACAAAACCTGACAACTCAAGGATTAAGAGAGATTTTAC 1414  
 QY 1494 ACTTTTATCACTCACTGCACTGATCGATCAAGAAATCAAGCTGTGAGATCTTCAG 1553  
 DB 1415 TCTGATGCCAATAAACCCTCAATAAGTCCAAAGTACTTGGGTTACCTTATCTTTTC 1474  
 QY 1554 AGACAGCTGTTCCACAGTACTGCTGCAAGCCCTATGAGGAGAGGCTTAAGTCAA 1613  
 DB 1475 TAACAAGCAAGTGAATAATCTTTCTTAAGACCTTTGGGACCTCATGATTAATTTCAA 1534  
 QY 1614 GTCACTGCACTGAATGAGCCAGCCCTTATGATGTTGAGACAGGAGACCTTCCAGAAAT 1673  
 DB 1535 ATCTGTCACTCAATGATGCTTAACTTAAAGATGATGATCAAACTTGTCCACCTTT 1594  
 QY 1674 GAAGCCCGCCCTTCCGAGCGGCGGACCAATGTCATCCCTCAGTCAACATGAGCTT 1733  
 DB 1595 AATGAAAAACCTCTCCGCGCAGAGAGTTCACTGGGCTGTGCACTTTCTCATATAGTTT 1654  
 QY 1734 TTATGTGTCAAGATGTCATGCTTTGGCTGTGCGCTACCGATTA 1779  
 DB 1655 TTTTGTGATGAATAATGCAAGTGTGCTGTGATCTGAATAATA 1700

RESULT 6  
 US-09-071-739B-1  
 Sequence 1, Application US/09071739B  
 Patent No. 6177345  
 GENERAL INFORMATION:  
 APPLICANT: Iris Pecker et al.  
 TITLE OF INVENTION: HEPARANASE SPECIFIC MOLECULAR PROBES  
 TITLE OF INVENTION: AND THEIR USE IN RESEARCH AND MEDICAL







MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
 COMPUTER: Twinhead\* 511mote-890TX  
 OPERATING SYSTEM: MS DOS version 6.2,  
 Windows version 3.11  
 SOFTWARE: Word for Windows version 2.0 converted to  
 an ASCII file  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/260,038B  
 FILING DATE: 02-Mar-1999  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 09/071,618  
 FILING DATE: May 1, 1998  
 APPLICATION NUMBER: 09/071,739  
 FILING DATE: May 1, 1998  
 APPLICATION NUMBER: 08/922,180  
 FILING DATE: September 2, 1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Friedmam, Mark M.  
 REGISTRATION NUMBER: 33,883  
 REFERENCE/DOCKET NUMBER: 910/16  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 972-3-562553  
 TELEFAX: 972-3-562554  
 TELEX: <Unknown>  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1721  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-09-260-038B-1

Query Match 15.7%; Score 278.8; DB 4; Length 1721;  
 Best Local Similarity 52.4%; Pred. No. 3.8e-76;  
 Matches 684; Conservative 0; Mismatches 592; Indels 30; Gaps 2;

```

QY 474 AGATTAACGAAAGGCTGCAAGTTGCCAGACCTGATGTTATGCTGCTCCAAAG 533
DB 425 AGTCAACGAGATTTGGCAAAATGATGATCCCTCTGATGAGGAGAAAGTTGCG 484
QY 534 GAGAGAGAGAGCTCAGATGATCTGCTCTTCTTAAAGAGCAATTCCTCAATCTTAAG 593
DB 485 GTTGAATGGCCCTACAGAGAGAAATGCTACTCGAGAGACTACAGAAAAATTGCA 544
QY 594 TATCTCATATTACACGCAAGTCTCTAGACAACTTTTAACTTTGCTGATGCTGCG 653
DB 545 GAACAGACCTTACTCAAGAGGCTGTAGATGCTATACCTTTTGCAACCTGCTCAGG 604
QY 654 ACTCCAGCTGATATTGCTCTAATGACCTGCGGTGATCCCAATTACTCTGGAACG 713
DB 605 ACTGAGCTTGATCTTGGCTTAATGCTTATTAGAACAGCAATTTGCACTGGAACG 664
QY 714 TTCTAGTCCCTGAGTCTGTGAAGTACAGCCGCAAGCAAAAAGTACAACTTTCTTGGCA 773
DB 665 TTCTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 724
QY 774 ACTGGTATGAGCCCAATTAATCATGAGACATGATGAGCGGAGAGTAAATGAGGACGCA 833
DB 725 ACTGAGCAATGACCTACACAGTTTCTTAAAGAGGCTGATATTTCATCAATAGGCTGCA 784
QY 834 GTTGGAGAAAGATTACATCCAGCTGAGAGGCTGTGAGCCCATCCGATTTATTCAG 893
DB 785 GTTAGAGAGAAAGTTATTAATTCATTAATGCAATAAC--TTCTAAGAAAGTCCACCTTCAAAA 841
QY 894 AGCAGCTATATGCGCTTAATATTTGGGGGCGCGAGAGAAATGTCATGCGCTCTGAGA 953
DB 842 TGCAAACTCTATGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 901
QY 954 TGGATTCATGAAGGTGAGAGAGTACAGTAAATGAGTACGTTACCTGCAACATTTGCTAAT 1013
DB 902 GAGCTTCTGAAAGGCTGAGAGAGTGAATGATGATGATGATGATGATGATGATGATGATGAT 961

```

```

QY 1014 TGATGCGCGGCTGCTCAAGGTATGAGACTTCTGAAAAGTCCGCTGTAGACACTCTC 1073
DB 962 GAATGAGCGAGCTGCTACAGAGAGATTTTCTAAACCTGATGATTTGGAATTTTAT 1021
QY 1074 TGACAGATTAGAAAATTCAGAAATGCTTATATACATCACTCCAGAGAAAGATTTTG 1133
DB 1022 TTATCTGCTCAAAAAGTTTCTCAGGTGTTGAGACCAAGGCTGCGCAAGAGGTCTG 1081
QY 1134 GCTTGAAGGTGTGTACCACTCAGCTGAGCAAAACATTCATTCGATTCCTATGCG 1193
DB 1082 GTTAGAGAAAACAAGCTCTGCATATGAGAGCGAGCGCTTGTCTATTCGACACTTTTCC 1141
QY 1194 TGACAGATTCTTATGCTTGAACAATTAGAAATGCTGCAATTCAGAGCAATGATGCTG 1253
DB 1142 AGCTGGTTTATGCTGCTGATTAATTTGGGCGCTGACGCCGCAATGGAAATGAGTGTAT 1201
QY 1254 GATAGCGCACTCATTTTGTGACCATGATACATCACTCTGTGACCAAGATTTTAAACC 1313
DB 1202 GATGAGCGCAAGTATTTCTTGGAGCAGAAACTACATTTAGTGAAGAAACTTGCATCC 1261
QY 1314 ATTACGAGACTACTGCTCTCTCTCTTACAAGCGCTGATGCGCCCAAGTCTTGCG 1373
DB 1262 TTATCTGATTTATGCTGCTATCTTCTTCTTCAAGAAATTTGGGCAACAGGTGTAT 1321
QY 1374 TGTGATGTGCTGGGCTCAGCGGAGCGCAAGCGCTGCGCAAGTATCCGGACAAACT 1433
DB 1322 GGCAGAGCTCAAGGTTCAAGAGAGAG-----AAGCT 1354
QY 1434 AAGGATTTATGCTCAGTGCAGCAAAACCAACCAACCACTAGTGTGCTGCTATAC 1493
DB 1355 TCGAGTATCTTATCTGCAACCAACCACTGCAAACTCAAGTATTAAGAGAGATTTAAC 1414
QY 1494 ACTTTTATCATCACTTGCATGATCAAGAAAGAAATCAAGTGGCTGGAAGTCTCAG 1553
DB 1415 TCTGTATGCCATTAACCTCATATACGTACCAAGTCTTGGCTTACCTTATCTTTTTC 1474
QY 1554 AGCAAGCTGCTTCAACAGTACTGCTGAGAGCCCTTGGGAGAGAGGCTTAAAGTCCAA 1613
DB 1475 TACAGCAAGTGTATTAATTAATCTTTTAAAGAGTGTGATGATCAACCTTGGCACTTT 1594
QY 1614 GTCAGTCAACTAATGAGCCAGCCCTTATGATGTGAGAGAGAGAGCCCTCCAGATTT 1673
DB 1535 ATCTGCTCACTCAATGATGTCTTAACTTAAAGATGTGATGATCAACCTTGGCACTTT 1594
QY 1674 GAAGCCCGGCGGCTTGGGCGGCGGAGCATTTGATGATCCCTGCAAGTGGGCTT 1733
DB 1595 AATGAGAAAACCTCTCGGCGAGAGATTTCACTGGGCTTGGCAAGCTTTCTCATATGATT 1654
QY 1734 TTATGTGTCAGAAATGTCATGCTTTGGCTGCGCTACCGATAA 1779
DB 1655 TTTTGATTAAGAAATGCCAAAGTGTGCTGCTGATCTGATAAATAA 1700

```

RESULT 9  
 US-09-260-038B-3  
 ; Sequence 3, Application us/09260038B  
 ; Patent No. 6348344  
 ; GENERAL INFORMATION:  
 APPLICANT: Maty Ayal-Herskovitz et al.  
 TITLE OF INVENTION: GENETICALLY MODIFIED CELLS AND METHODS FOR  
 EXPRESSING RECOMBINANT HEPARANASE  
 AND METHODS OF PURIFYING SAME  
 NUMBER OF SEQUENCES: 25  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Mark M. Friedman c/o Anthony Castorina  
 STREET: 2001 Jefferson Davis Highway, Suite 207  
 CITY: Arlington  
 STATE: Virginia  
 COUNTRY: United States of America  
 ZIP: 22202  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk

COMPUTER: Twinhead\* Slimnote-890TX  
 OPERATING SYSTEM: MS DOS version 6.21  
 SOFTWARE: Word for Windows version 3.11 converted to an ASCII file

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/260,038B  
 FILING DATE: 02-Mar-1999  
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 09/071,618  
 FILING DATE: May 1, 1998  
 APPLICATION NUMBER: 09/071,739  
 FILING DATE: May 1, 1998  
 APPLICATION NUMBER: 08/922,180  
 FILING DATE: September 2, 1997

ATTORNEY/AGENT INFORMATION:  
 NAME: Friedmann, Mark M.  
 REGISTRATION NUMBER: 33,883  
 REFERENCE/DOCKET NUMBER: 910/16

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 972-3-562553  
 TELEFAX: 972-3-562554  
 TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1721  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
 US-09-260-038B-3

Query Match 15.7%; Score 278.8; DB 4; Length 1721;  
 Best Local Similarity 52.4%; Pred. No. 3.8e-76;  
 Matches 684; Conservative 0; Mismatches 592; Indels 30; Gaps 2;

```

QY 474 AGATAAAGAAAGGTGCAAGTTGCCAGCAGCCTGTGATGTTTGTCTGTCTGCTCCAAAG 533
DB 425 AGCAACACAGAGATATTGCAAAATATGATTCATCTCTGATGTGAGAGAAAGTTGCG 484
QY 534 GGAGAGAGAGCTCAGATGCATCTGTTCTTCTAAAGAGCAATTCCTCAATCTTACAG 593
DB 485 GTTGAATGGCCCTACAGAGAGCAATGCTACTCCAGAAACCTACCAAAAAGTTCAA 544
QY 594 TAATCTCATTTTAAACGCGCAAGTCTCTAAGCAAACTTTTAACTTTGCTGATGCTGG 653
DB 545 GAACAGACCTTACTCAAGAGAGCTGTAGATGTCTATACCTTTTGCACAACTGCTCAGG 604
QY 654 ACTCCAGCTGATTTTGTCTCTAAATGCACTGGGTGTATCCCAATPACTCCTGGAACG 713
DB 605 ACTGACCTGATTTTGGCTTAATGCTTTATTAAGACAGAGATTGCAATGGAACAG 664
QY 714 TTCTAGTCCCTGAGTCTGTGAAGTACGCGCAACAAAAGTACCAATTTCTTGGGA 773
DB 665 TTCTATGCTCAGTGTCTCTGAGACTAGCTCTTCCAGGAGTATTAACATTTCTTGGGA 724
QY 774 ACTGGGTATGAGCCAAATTAAGTACGACCAATGATGCGCGGAGGAGTAAATGGAGCA 833
DB 725 ACTAGCAATGAACCTTAACGTTTCTTAAAGAGGTGATTTTTCATCATATGGGTGCA 784
QY 834 GTTGGAGAGATTAATCATCAGCTGAAGAGCTGTGACAGCCCATCCGATTTATTTCCAG 893
DB 785 GTTAGAGAGATTAATTAATTCATGATTAAC---TTTAAAGAGTCAACCTTCAAAA 841
QY 894 AGCCAGCTTAATGGCCCTAATATTTGGGGCGGAGAGAAATGTCATCCCTCTTACA 953
DB 842 TGCATACTCTATGCTGATGTGTGTAAGCTCAGAGAAAGACGGCTTAAGATGTGA 901
QY 954 TGGATTCATGAAGGTGAGAGAGTACAGTAGATGAGTTACCTGGCAACATTTGCTACAT 1013
DB 902 GAGCTTCGTAAGGCTGTGAGAGAGTGAATGATTCAGTTACATGGCATCACTACTATT 961

```

```

QY 1014 TGATGCGCGGTGTGTCAGGTGATGACTTCTGAAAACCTGGCTGTAGACACTCTC 1073
DB 962 GATGAGCGAGCTGTCACAGGAAATTTTCTTAAACCTGATGTATTTGACATTTTAT 1021
QY 1074 TGACCAATTAGAGAAATTCAGAAAGTGTAAATACATACCTCCAGAAAGATTTG 1133
DB 1022 TTCACTGTGCAAAAAGTTTCCAGGGGTGTGAGACACAGGCTGGCAAGAGGTCTG 1081
QY 1134 GCTTGAAGGTGTGTGACCACTCAGTGTGAGGCAAAAACATCTATCCGATTCCTATG 1193
DB 1082 GTTAGAGAAACAACTCTGCAATGAGAGGAGGCGCTTGTCTATCCAGCTTTTC 1141
QY 1194 TGCAGATTCCTTATGTTGAACCTTTAGGAATGCTGGCCCAATCAGGCAATGATGCT 1253
DB 1142 AGCTGGCTTATGAGCTGTGATTAATTTGGCTGTGACGCCAATGGAAATAGAGTGTG 1201
QY 1254 GATACGGCACTATTTTTCACATGATATCATCAGCTGTGACCAAGATTTTAACC 1313
DB 1202 GATAGGCAAGTATTTCTTGAAGAGAACTACATTTAGTGAAGAACTTCATCCATCC 1261
QY 1314 ATTACGAGACTAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1373
DB 1262 TTACCTGATTTATGCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1321
QY 1374 TGTGATGTGCTGGCTCTCCAGCGGAGCCAGGCTGCGCGAGTATCCGGACAACT 1433
DB 1322 GGCAGAGGTGTGAAAGTTCAAAAGAGAG-----AAGCT 1354
QY 1434 AAGATTTATGCTCACTGACACAAACCAACCAACCACTAGTTGTTGATTCATTC 1493
DB 1355 TCGAGTATCTCTTATGCTACAAACCTGACCAATCCAGGTATTAAGAGAGATTTAC 1414
QY 1494 ACTTTTATCATCACTTGTGATGATCAGAAAGAAATCAAGCTGTGCTGGACTCTCAG 1553
DB 1415 TCTGTATGCAATAAACCTCCATATACGTCACCAAGTCTTGGGTATCCCTATCTTTTC 1474
QY 1554 AGACAGCTGTTTACAGTACCTGCTGAGCGCTTATGGGACAGAGGCTTAAAGTCCA 1613
DB 1475 TAAAGAGAAAGTAAATTAATCTTTTAAAGCTTTGGAGCTCAGATTAATCTTCCA 1534
QY 1614 GTCAGTCAACTGATAGGCGAGCCCTTATGATGTGAGAGAGAGCCCTCCAGAAAT 1673
DB 1535 ATCTGTCACTCATGATGTCTTAACTTAAAGATGTGATCAAACTTGCACCTTT 1594
QY 1674 GAACCCCCCGCTTGGGCGCGCGGAGATTTGTTCATCCCTCCAGTACATGGCTT 1733
DB 1595 AATGAAAAACCTCTCGGCGAGAGATTCATGCGGCTTGCAGCTTTCTCATATGTTT 1654
QY 1734 TTATGTGTCAGAAATGTCAATGCTTTGGCTGCGCTCCGCTACCGATTA 1779
DB 1655 TTTTGATTAAGAAATGCCAAAGTGTCTGTGATCATGAAATTA 1700

```

RESULT 10  
 US-09-635-923-1  
 ; Sequence 1, Application US/09635923  
 ; Patent No. 6426209

GENERAL INFORMATION:  
 APPLICANT: Maty Ayal-Hershkovitz et al.  
 TITLE OF INVENTION: GENETICALLY MODIFIED CELLS AND METHODS FOR  
 EXPRESSING RECOMBINANT HEPARANASE  
 AND METHODS OF PURIFYING SAME

NUMBER OF SEQUENCES: 25  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Mark M. Friedman c/o Anthony Castorina  
 STREET: 2001 Jefferson Davis Highway, Suite 207  
 CITY: Arlington  
 STATE: Virginia  
 COUNTRY: United States of America  
 ZIP: 22202

COMPUTER READABLE FORM:  
 MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
 COMPUTER: PC

```

OPERATING SYSTEM: MS DOS version 6.2,
SOFTWARE: Word for Windows version 3.11
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/635,923
  FILING DATE: 10-Aug-2000
  CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US/09/487,716
  FILING DATE: 19-Jan-2000
  APPLICATION NUMBER: 09/071,618
  FILING DATE: May 1, 1998
  APPLICATION NUMBER: 09/071,739
  FILING DATE: May 1, 1998
  APPLICATION NUMBER: 08/922,180
  FILING DATE: September 2, 1997
ATTORNEY/AGENT INFORMATION:
  NAME: Friedmann, Mark M.
  REGISTRATION NUMBER: 33,883
  REFERENCE/DOCKET NUMBER: 910/16
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 972-3-5625554
  TELEFAX: 972-3-5625554
  TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 1:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 1721
    TYPE: nucleic acid
    STRANDEDNESS: double
    TOPOLOGY: linear
  SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-635-923-1

Query Match      15.7%; Score 278.8; DB 4; Length 1721;
Best Local Similarity 52.4%; Pred. No. 3.8e-76;
Matches 684; Conservative 0; Mismatches 592; Indels 30; Gaps 2;

```

```

OY 1014 TGATGCGCGGTGTGTGAGGTGAGTCTTGAAAACCTGGCTGTGACACACTCTC 1073
DB 962 GAATGACGCGACTGTACACAGGAGAAATTTTCTAAACCTGTATGTGACATTTTAT 1021
OY 1074 TGACAGATTAAGGAAATTCAGAAAGGTATATACATCACTCCGAGAAAGATTG 1133
DB 1022 TTGATCTGTCAAAAATTTCAGAGGTGTGAGACACAGGCTCGGCAAGAGTCTG 1081
OY 1134 GCTTGAAGGTGTGTGACCACTCAGCTGAGGCAAAAATATATCCGATCTTATGC 1193
DB 1082 GTTAGAGAAACAAGCTCTGTATGAGAGGCGAGCGCTTGTCTATCCGACACTTTCG 1141
OY 1194 TGAGAGATTTTATGTGTGAACCTTAGAAGATCTGGCCAAATCAGAGGATGATGTGT 1253
DB 1142 AGCTGGCTTTATGTGTGTGATTAATTTGGCTGTGAGCCGCAATGGAATGAGATG 1201
OY 1254 GATACGCGACTCATTTTGTGACCATGATATCAATCACTGTGTGACAGAAATTTTAA 1313
DB 1202 GATGAGCAAGTATTTCTTTGAGACAGAAATCAATTTAGTGTGAAAACTTCATCC 1261
OY 1314 ATTTCCAGTACTGTGCTCTCTCTCTCTCAAGCGCTGATGTGCGCCAAAGTCTTGC 1373
DB 1262 TTAACTGATTTATGTGTATCTTCTGTCAAGAAATTTGTGGGACCAAGGTGTTAT 1321
OY 1374 TGTGATGTGTGCTGGCTCCAGCGGAGGACAGGCTGTGCGGAGTATCCGGACAACT 1433
DB 1322 GCGAGGTGTCAAGGTTCAAAGGAAG-----AAGCT 1354
OY 1434 AAGATTTATGTCTACTGTGCAAAACCAACCAACCAACTACCTTGTGTGCTTAC 1493
DB 1355 TCGAGTATACCTTCATTTGCAAAACCTGACATTCAGGTATTAAGAGGAGATTTAC 1414
OY 1494 ACTTTTATCATCACTGTGATGATCAAGAAAGAAATCAAGCTGTGTGAGCTCTGAG 1553
DB 1415 TCTGTATGCATTAACCTCATATAGCTACAGATCTTGGCTTACCTTATCTTTTC 1474
OY 1554 AGACAGCTGTGTACACAGTACTGTGTGAGCCCTTATGTGGGAGAGGAGGCTTAA 1613
DB 1475 TAACAGCAAGTGTATTAATTAATCTTAAAGCTTTGGAGACCTTATGATTAATCTT 1534
OY 1614 GTGAGGCACTGAATGTGAGGCTTGAATGTGATGATGATGATGATGATGATGATG 1673
DB 1535 ATCTGTCAACTCAATGTGTAACTCTTAAGATGTGATGATGATGATGATGATGAT 1594
OY 1674 GAAGCCCCGCCCCCTTGGGCGGCGGACATTTGATCTCTCCAGTACCATGGGCTT 1733
DB 1595 AATGAAAAACCTCTCGGCGAGAAATTCATCGGCTTGCAGCTTCTCATATAGT 1654
OY 1734 TTATGTGTGAAGATGTCAATCTTTGGCTGGCCCTTACCATTA 1779
DB 1655 TTTGTGATTAAGAAATGCCAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1700

RESULT 11
US-09-635-923-3
Sequence 3, Application US/09635923
Patent No. 6426209
GENERAL INFORMATION:
APPLICANT: Mary Ayal-Hershkovitz et al.
TITLE OF INVENTION: GENETICALLY MODIFIED CELLS AND METHODS FOR
EXPRESSING RECOMBINANT HEPARANASE
AND METHODS OF PURIFYING SAME
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Mark M. Friedmann c/o Anthony Castorina
STREET: 2001 Jefferson Davis Highway, Suite 207
CITY: Arlington
STATE: Virginia
COUNTRY: United States of America
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
COMPUTER: PC

```

OPERATING SYSTEM: MS DOS version 6.2,  
Windows version 3.11  
SOFTWARE: Word for Windows version 2.0 converted to an ASCII file  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/635,923  
FILING DATE: 10-Aug-2000  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/487,716  
FILING DATE: 19-Jan-2000  
APPLICATION NUMBER: 09/071,618  
FILING DATE: May 1, 1998  
APPLICATION NUMBER: 09/071,739  
FILING DATE: May 1, 1998  
APPLICATION NUMBER: 08/922,180  
FILING DATE: September 2, 1997

ATTORNEY/AGENT INFORMATION:  
NAME: Friedman, Mark M.  
REGISTRATION NUMBER: 33,883  
REFERENCE/DOCKET NUMBER: 910/16

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 972-3-562553  
TELEFAX: 972-3-562554

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1721  
TYPE: nucleic acid  
STRANDEDNESS: Double  
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-635-923-3

Query Match 15.7%; Score 278.8; DB 4; Length 1721;  
Best Local Similarity 52.4%; Fred. No. 3.8e-76;  
Matches 684; Conservative 0; Mismatches 592; Indels 30; Gaps 2;

```

QY 474 AGATTAACAGAAAGGCTGCAAGATTGCCAGCACCCCTGATGTTATGCTGGTCCCAAG 533
DB 425 AGTCAACAGAGATTTGCAATATGATCCATCCCTCTGATGAGAGAAATTACG 484
QY 534 GAGAAGGACCTCAGATCATCTGTTCTTAAAGAGCAATTCACATCTTACAG 593
DB 485 GTTGAATGCCCCACAGAGACATGCTACTCCAGAAACATACAGAAAGTTCAA 544
QY 594 TAACTCATTAACAGCAGCTCTTACAGAACTTTATACCTTGTCTGATGCTGG 653
DB 545 GAAACAGACCTACTCAAGAAAGCTGTAGATGCTATACCTTTTCAAACTGCTCAG 604
QY 654 ACTCACCTGATATTTGCTTAAAGCACTGCTGTATCCCAATTAACCTCTGGAGAG 713
DB 605 ACTGACCTGATCTTTGGCCTTAATGCTTATTAAGAACAGCAATTTGCACTGAGAG 664
QY 714 TTCTAGTCCCTGAGTGTGTTGAAGTACAGGCGCAAAAAAGTAAACATTTCTGGGA 773
DB 665 TTCTAATGCTCAGTTGCTCTGAGACTACTGCTCTTCAAGAGGATTAACATTTCTGGGA 724
QY 774 ACTGGTATAGGCAATAATACATACGACATGAGCAGGCGGCACTTAATAGGCAAGCA 833
DB 725 ACTAGGCAATACCTTAACAGCTTCTTAAAGAGCTGATATTTTCACTCAATGAGTGGCA 784
QY 834 GTTGGAAAGGATTACATCCAGCTGAAGAGCTGTGCAAGCCATCCGGAATTATTCAG 893
DB 785 GTTGGAGAAAGATTATATTCATTGCAATAAC--TTCTAAGAAAGTCCACCTTCAAAA 841
QY 894 AGCAGCTTATATGAGCCCTATATTTGGCGGCGGAGAGAAAGATGCAATCCGCCCTCAGA 953
DB 842 TGCAAACTCTATGCTCTGATGTTGGTCAAGCTCGAAGAAAGCGCTAAGATGCTGAA 901
QY 954 TGAATTATGAGGTGCGAGAAATACAGTATGATGCAAGTTACCTGGCAACATTTGCTCAT 1013
DB 902 GAGCTCTCTGAAGGCTGGTGAAGAAAGTATGATTGATTACATGACATCACTATATT 961

```

```

QY 1014 TGATGCGCGGATGTCAAGTGTATGACTCTCTGAAAACTGCCCTGTAGACACACTCTC 1073
DB 962 GAATGAGGAGACTGTACACAGGAAGATTTTCTAAACCTATGATTGACATTTAT 1021
QY 1074 TGACCAATTTAGAAAATTACAAAAGTGTATATACACTCCAGAAAGAGATTGG 1133
DB 1022 TTCACTGTGCAAAAAGTTTCCAGGTGTTAGAGACACAGGCTGCAAGAAAGTGTG 1081
QY 1134 GCTTGAAGGTGTGTGACCACTCAGCTGAGAGCAAAACATCTATCCGATCTATAGC 1193
DB 1082 GTTAGAGAAAACAGCTGTGATATGAGGCGAGGCGCTTGTATCCACCTTTGCG 1141
QY 1194 TCGAGATTTATGATTTGAACACTTTAGAAATGCTGCAATCAAGGCAATGATGCTG 1253
DB 1142 ACTGCTCTTATATGCTGTGATTAATTTGGGCTGTCAAGCCGAATGGAAATGAAGTGT 1201
QY 1254 GATAGGCACTGATTTTGTGACATGATACACTGCTGTGACGCAAAATTTAAACC 1313
DB 1202 GATGAGGCAAGTATCTTTGAGAGCAAGAACTACCATTTAGTGAATGAATCTTGATCC 1261
QY 1314 ATTACGAGACTACTGAGCTCTCTCTCTCTACAGGCGCTGATCGGCCCCAAAGTCTTGGC 1373
DB 1262 TTACCTGATTAATGGCTATCTCTTCTGTAAGAAATGGTGGCACCAAGTGTATAT 1321
QY 1374 TGTGATGTGCTGGGCTCCAGCGGACAGCCGCTGCGAGTATCCGGACAACT 1433
DB 1322 GGCAGCGGTGCAAGGTTCAAAAGAAAG--AGCT 1354
QY 1434 AAGGATTTATGCTCACTGCAAAACCAACACCACTACGTTCTGGGTCCATTAAC 1493
DB 1355 TCGAGTATACCTTCAATGTGCAAAACAGTCAATCAAGTATTAAGAAAGATTTAAC 1414
QY 1494 ACTTTTATCATCACTTGCATCGATCAAGAAAGAAATCAAAGTGGCTGGACCTCTCAG 1553
DB 1415 TCTGATGCAATAAACCCTCATACGTCACCAAGTACTGCGGTTACCTTCTTTC 1474
QY 1554 AGCAAGCTGTTCACAGTACCTGCTGAGCCCTATGAGGCGAGAGGCTTAAAGTCAA 1613
DB 1475 TAAACAGCAAGTGAATTAATTAATCTTCTTAAGACTTTGGACCTCATGATTAATTTCAA 1534
QY 1614 GTCACTGCAAGTGAATGCGACGCTTATGATGTTGAGCAAGAGGACCTTCCAGAAAT 1673
DB 1535 ATCTGTCAACTCAATGATGCTTACTTAAGATGTTGATGATCAAACTTCCACCTTT 1594
QY 1674 GAAAGCCCCCCTTCCGCGCGGCGGAGCAATGTCATCCCTCGAGTCAACCATGGGCTT 1733
DB 1595 AATGAAAAACCTTCCGCGCGAGAGTTCACTGGGCTTGCAGCTTTCTCATATAGTTT 1654
QY 1734 TATATGTCAGAAATGTCATGCTTTGGCCTGCGCTAACGATTA 1779
DB 1655 TTTTGTGATTAAGAAATGCCAAGTTGCTGCTTGCATCGAAATTA 1700

```

RESULT 12  
US-09-487-716A-1  
Sequence 1, Application US/09487716A  
Patent No. 6475763  
GENERAL INFORMATION:  
APPLICANT: Mary Ayal-Hershekovitz et al.  
TITLE OF INVENTION: GENETICALLY MODIFIED CELLS AND METHODS FOR  
EXPRESSING RECOMBINANT HEPARANASE  
AND METHODS OF PURIFYING SAME  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Mark M. Friedman c/o Anthony Castorina  
STREET: 2001 Jefferson Davis Highway, Suite 207  
CITY: Arlington  
STATE: Virginia  
COUNTRY: United States of America  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
COMPUTER: PC

Fri May 14 13:01:53 2004

us-10-088-676-1.rml

Page 12

OPERATING SYSTEM: MS DOS version 6.2,  
Windows version 3.11  
SOFTWARE: Word for Windows version 2.0 converted to an ASCII file  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/487,716A  
FILING DATE: 19-Jan-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/071,618  
FILING DATE: May 1, 1998  
APPLICATION NUMBER: 09/071,739  
FILING DATE: May 1, 1998  
APPLICATION NUMBER: 08/922,180  
FILING DATE: September 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Friedman, Mark M.  
REGISTRATION NUMBER: 33,883  
REFERENCE/DOCKET NUMBER: 910/16  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 972-3-562553  
TELEFAX: 972-3-562554  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1721  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-487-716A-1  
Query Match 15.7%; Score 278.8; DB 4; Length 1721;  
Best Local Similarity 52.4%; Pred. No. 3.8e-76;  
Matches 684; Conservative 0; Mismatches 592; Indels 30; Gaps 2;  
QY 474 AGATTAACAGAAAGCTGCAAGATTGCCAGACCCCTGATGTTATGCTGCTCCAAAG 533  
DB 425 AGTAAACAGAGATTTTGGCAAAATATGATCCATCCCTCTGATGTCAGAGAACTTTCG 484  
QY 534 GAGAAAGGCACTCAGATGATGCTGCTTCTTAAAGAGCAATCTCAATATTCTTACAG 593  
DB 485 GTTGAGATGGCCCTACAGAGCAATTTGCTACCTCCAGAACTACACGAAATTTTAA 544  
QY 594 TAACTCATTAACAGCCAGCTCTCTAAGCAAACTTTTAACTTCTGATGCTGCTG 653  
DB 545 GAAAGAGCACTACTCAAGAAAGCTCTGATGATGCTATACCTTTTGAACATGCTCAG 604  
QY 654 ACTCAGCTGATATTTCTCTAATGCACTGCTGCTAATCCCAATACCTCTGAGAG 713  
DB 605 ACTGAGCTGATGCTTGGCTTAATGCTTATTAAGACAGAGATTTGCACTGAG 664  
QY 714 TTCTAGTCCCTGAGTCTTGTGAAGTACAGCCGACAGAAAGTACACATTTCTTGG 773  
DB 665 TTCTAAGTCTGATGCTCTGAGCTAATGCTCTTCCAAAGGGGTAAACATTTCTTGG 724  
QY 774 ACTGCTTAATGAGCAAAATTAATGAGCAATGAGCAATGAGCCGAGAGTAAATGAG 833  
DB 725 ACTAGGAGATGAACCTTAACATTTCTTAAGAGGCTGATATTTTCAATAGGCTGCA 784  
QY 834 GTTGGGAAAGATTAATCACTCAGCTGAAGAGCTGTTGAGCCCAATCCGATTTATCC 893  
DB 785 GTTAGGAGAAATTAATTAATGCTTAAC---TTCTAAGAAAGTCCACCTTCAAAA 841  
QY 894 AGCAGCTTAATGAGCCCTTAATTTGGGCGGCGAGAGAAATGTCATGCGCTCTAGA 953  
DB 842 TGCAAACTCTATGCTGCTGATGCTGCTGCTCCAGAAAGAGCGCTTAAGATCTGA 901  
QY 954 TGATTCATGAGGCTGAGCAAGAGTACAGATGAGTACAGTACCTGAGCAATGTTACAT 1013  
DB 902 GAGCTTCTGAGAGCTGCTGAGAGATGATGATTTGATTAATGAGCACTAATTAATTT 961  
QY 1014 TGATGCGGCTGCTCAGGCTGATGATCTTCTGAAAACCTCGCTGTTAGACACCTCTC 1073

DB 962 GATGAGCGAGCTGCTACAGAGAAATTTCTTAAACCTGATGATTTGACATTTTAT 1021  
QY 1074 TACCAAGATTGAGAAATTTGAGAAAGTGTAAATACATCACTCCAGAAAGATTTTG 1133  
DB 1022 TTCACTCTGAGAAAGATTTTCCAGGCTGTTAGAGACACAGGCTTGGCAAGAGCTTG 1081  
QY 1134 GCTTAAGAGTGTGAGACCACTCACTGAGAGGCAAAACATCTATCCGATTTCTATCC 1193  
DB 1082 GTTAGAGAAACAGCTCTGATATGAGAGGCGCCCTTGTATCCGACCTTTGCT 1141  
QY 1194 TCCAGATTTCTTAATGTTGAACACTTGAAGATGCTGCGCAATCAGGCAATGATGCT 1253  
DB 1142 AGCTGCTTTATGCTGCTGATTAATTTGGCTGCTGCTGAGCCCAATGGAATGAAATG 1201  
QY 1254 GATAGGAGCTCATTTTGTGACATGATACATCACTCTGAGAGCAAAATTTTAAACC 1313  
DB 1202 GATGAGGAGATTTCTTTGAGAGAGAACTACATTTTGTGATGAAACTTTCATCC 1261  
QY 1314 ATTACAGACTACTGCTCTCTCTCTAACAAGGCTGATGAGCCCAAGCTTGGC 1373  
DB 1262 TTACCTGATTTATGCTATCTCTCTCTGCTCAAGAAATGCTGAGCAAGCTGTTAT 1321  
QY 1374 TGTGATGCTGCTGCTGCTCCAGAGGAGCAAGGCTGCGCAATGATCCGAGCAACT 1433  
DB 1322 GGCAGCTGCAAGCTTCAAGAGAGG-----AGCT 1354  
QY 1434 AAGATTTATGCTCACTGCAACAAACCAACCAACCACTACCTGCTGCTCATTC 1493  
DB 1355 TCGATATACCTTCACTGCAACAAACCACTGCAACCACTGCAATTCAGGATTAAGAGATTTAAC 1414  
QY 1494 ACTTTTATCACTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1553  
DB 1415 TCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1474  
QY 1554 AAGCAAGCTGTTTCAAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1613  
DB 1475 TAAACAGAGATGATTAATATCTTCTTGAAGCTTTGGAGCTTCAATGATTTTCCAA 1534  
QY 1614 GTCACTGCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1673  
DB 1535 ATCTGCTCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1594  
QY 1674 GAGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1733  
DB 1595 AATGAGAAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1654  
QY 1734 TTATGCTCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1779  
DB 1655 TTTTGTGATTAAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1700  
RESULT 13  
US-09-487-716A-3  
Sequence 3, Application US/09487716A  
Patent No. 6475763  
GENERAL INFORMATION:  
APPLICANT: Maty Ayal-Hershkovitz et al.  
TITLE OF INVENTION: GENETICALLY MODIFIED CELLS AND METHODS FOR  
EXPRESSING RECOMBINANT HEPATINASE  
AND METHODS OF PURIFYING SAME  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Mark M. Friedman c/o Anthony Castorina  
STREET: 2001 Jefferson Davis Highway, Suite 207  
CITY: Arlington  
STATE: Virginia  
COUNTRY: United States of America  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
COMPUTER: PC  
OPERATING SYSTEM: MS DOS version 6.2,  
Windows version 3.11



```

SOFTWARE: Word for Windows version 2.0 converted to an ASCII file
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/487,716A
FILING DATE: 19-Jan-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/071,618
FILING DATE: May 1, 1998
APPLICATION NUMBER: 09/071,739
FILING DATE: May 1, 1998
APPLICATION NUMBER: 08/822,180
FILING DATE: September 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Friedman, Mark M.
REGISTRATION NUMBER: 33,883
REFERENCE/DOCKET NUMBER: 910/16
TELECOMMUNICATION INFORMATION:
TELEPHONE: 972-3-562553
TELEFAX: 972-3-562554
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1721
TYPE: nucleic acid
STRANDEDNESS: Double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-487-716A-3

Query Match      15.7%; Score 278.8; DB 4; Length 1721;
Best Local Similarity 52.4%; Pred. No. 3.8e-76;
Matches 684; Conservative 0; Mismatches 592; Indels 30; Gaps 2;

474 AGATAAACAGAAAGCTGCAAGATTGCCAGACCTGATGTTATGCTGTCCTCAAG 533
425 AGTCAACAGAGATTTTGCAAAATGATCCATCCCTCGATGTGAGAGAGATTACG 484
534 GGAAGAGAGCTAGATGATCTGTTCTTCTAAAGAGCAATTCCTCAATCTTACG 533
485 GTTGAATGGCCCTACAGAGCAATGCTACTCCGAGAACACTACAGAAAGTTTCA 544
594 TAATCTCATATTAAACAGCAGCTCTAGACAACTTATTAATCTTGAATGCTGCG 653
545 GAAACAGACCTACTCAAGAACTCTGATAGATGCTATACCTTTGCAAACTGCTCAG 604
654 ACTCCACTGATTTTGTCTTAATGCACTGCTGCTGATCCCAATATCTCGAAGCAG 713
605 ACTGCACTGATCTTGGCTTAATGCTTAATTAAGAACAGCAATTTGCAATGAGCAG 664
714 TTCTAGTCCCTGAGTCTGTTGAAGTACAGCGCCAGCAAAAAGTACAACTTTCTGGA 773
665 TTCTAATCTGATGCTGCTCCCTGCACTACTGCTCTTCCAGGGGTAAACATTTCTTGGGA 724
774 ACTGGGTAATAGCCCAATAAATCTATCGGACCATGATGCGGCGAGTAATGCGACCA 833
725 ACTAGGCAATGAACTTAAACATTTCTTAAAGAGGCTGATATTTTCAATGAGGTCGCA 784
834 GTTGGGAAAGGATTAATCCAGCTGAGAGGCTGTGAGCCCATCGGAGTTTATTCAG 893
785 GTTAGAGAGATTAATTAATTCATGCAATTAAC--TTCTAAGAAAGTCCACCTTCAAAA 841
894 AGCCAGCTTATATGCGCTTAATATTTGGCGCGCCGAGAGAAAGTATCAATGCGCTCTAGA 953
842 TGCAAAACTATATGCTGCTGATGTTGTCAGCTCGAAGAAAGACGCTAAGATGCGAA 901
954 TGAATTCATGAAGTGGAGAGAGTACAGTATGCAATGCGCTTCTGGAAGCTTCTCAT 1013
902 GAGCTTCTGAAGGCTGGTGAAGATGATGATTCATGATGCAATGCACTACTATTT 961
1014 TGATGGCCGGGTGTCAAGTATGATGACTTCTGTAACCTGCTGTTAGACACACTCTC 1073
962 GAATGACGAGACTGTCTACAGGAAAGATTTCTTAAACCTGATGATTTGCAATTTTAT 1021

```

```

1074 TGACCAATTAAGAAATTCAGAAAGTGTATATACATACCTCCAGAAAGAAATTTG 1133
1022 TTCACTGTGCAAAAAGTTTTCAGAGTGTGAGACACAGGCTGCGCAAGAGCTTG 1081
1134 GCTTGAAGTGTGTGACCACTCAGCTGAGGCAAAACATCTATCCGATTCCTATG 1193
1082 GTTAGAGAAACAGCTCTGATATGAGGGGAGCGCCCTTGCTGATCGACCTTTGC 1141
1194 TCCAGATTTCTATGTTGAACACTTTAGATGCTGCGCAATCAGGCAATGATGCTG 1253
1142 ACTGCTTTATGTTGCTGATTAATGAGGCTGTCAGCCGGAATGGAATGAGTGT 1201
1254 GATAGGCACTGATTTTGTGACATGATTAATTAATCTGCTGACCAATTTTAAACC 1313
1202 GATGAGGCAAGTATCTTTGAGCAGAACTACCAATTAATGATGAAATCTGATCC 1261
1314 ATTACAGACTAGCTGCTCTCTCTCTACAGAGGCTGATGCGCCCAAGCTTTGCG 1373
1262 TTTAAGTATTAATGCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1321
1374 TGTGATGTGCTGAGCTCCAGCGGAGCAAGGCTGCGGAGTATCCGAGCAAACT 1433
1322 GCAAGCGTGAAGGTTCAAGAGAGAG-----AGCT 1354
1434 AAGATTTATGCTCATCTGCAACAAACCAACCAACCAACCAACCAACCAACCAAC 1493
1355 TCGAGTATACCTTCAATGCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1414
1494 ACTTTTATCATCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1553
1415 TCTGATGCAATTAACCTTCATTAACCTTCATTAACCTTCATTAACCTTCATTAAC 1474
1554 AAGCAAGCTGTTTCAACGATCTGCTGATGATGATGATGATGATGATGATGATGAT 1613
1475 TAACAAGAGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1534
1614 GTCAAGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1673
1535 ATCTGCACTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1594
1674 GAAAGCCGCGCCCTTCCGCGCGCGCGGACATGATGATGATGATGATGATGATGAT 1733
1595 AATGAAAAAATCTTCCGCGCGGAGATGATGATGATGATGATGATGATGATGATGAT 1654
1734 TTAATGATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1779
1655 TTTTGTATGAAGAAATTCAGAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1700

RESULT 14
US-09-322-977-1
Sequence 1, Application US/09322977.
Patent No. 6531129
GENERAL INFORMATION:
APPLICANT: ILS Becker et al.
TITLE OF INVENTION: HEPARANASE SPECIFIC MOLECULAR PROBES
TITLE OF INVENTION: AND THEIR USE IN RESEARCH AND MEDICAL
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Mark M. Friedman c/o Anthony Castorina
STREET: 2001 Jefferson Davis Highway, Suite 207
CITY: Arlington
STATE: Virginia
COUNTRY: United States of America
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
COMPUTER: Windows 95/NT
OPERATING SYSTEM: MS DOS version 6.2,
SOFTWARE: Word for Windows version 2.0 converted to
CURRENT APPLICATION DATA:

```



APPLICATION NUMBER: US/09/322,977  
 FILING DATE: 08/922,180  
 CLASSIFICATION: 08/922,180  
 PRIOR APPLICATION NUMBER: 08/922,180  
 FILING DATE: September 2, 1997  
 APPLICATION NUMBER: 09/071,739  
 FILING DATE: May 1, 1998  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Friedman, Mark M.  
 REGISTRATION NUMBER: 33,883  
 REFERENCE/DOCKET NUMBER: 910/21  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 972-3-562553  
 TELEFAX: 972-3-562554  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1721  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 US-09-322-977-1

Query Match 15.78; Score 278.8; DB 4; Length 1721;  
 Best Local Similarity 52.4%; Pred. No. 3.8e-76;  
 Matches 684; Conservative 0; Mismatches 592; Indels 30; Gaps 2;

474 AGATTAACAAAGAGGCTGCAAGATGCGCCAGCACCCCTGATGATGCTGCTCAAG 533  
 425 AGTCAACAGAGATATTTGCAATATGATCCATCCCTCTGATGAGGAGAGATTAC 484  
 534 GAGAGAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 593  
 485 GTTGAGATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 544  
 594 TTAATCATATTAACACCCAGCTGCTAGACAACTTTATATCTTGATGATGATGAT 653  
 545 GAAAGAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 604  
 654 ACTGACCTGATATTTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 713  
 605 ACTGACCTGATATTTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 664  
 714 TTCTAGTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 773  
 665 TTCTAGTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 724  
 774 ACTGATATGAGGCAATATATCTGATGATGATGATGATGATGATGATGATGATGAT 833  
 725 ACTGATATGAGGCAATATATCTGATGATGATGATGATGATGATGATGATGATGAT 784  
 834 GTTGAGAGAGAGATTAATCACTGATGATGATGATGATGATGATGATGATGATGAT 893  
 785 GTTGAGAGAGAGATTAATCACTGATGATGATGATGATGATGATGATGATGATGAT 841  
 894 AGGAGCTTATATGAGGCAATATATCTGATGATGATGATGATGATGATGATGATGAT 953  
 842 TGCAGAACTCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 901  
 954 TGAATTCATGAGGTGAGGCAATATATCTGATGATGATGATGATGATGATGATGAT 1013  
 902 GAGCTTCTGAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 961  
 1014 TGATGAGGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1073  
 962 GATGAGAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1021  
 1074 TGACCAATATGAGAGAGATTAATCACTGATGATGATGATGATGATGATGATGAT 1133  
 1022 TTCTAGTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1081  
 1134 GCTTGAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1193

1082 GTTGAAGAGAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1141  
 1194 TGCAGATTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1253  
 1142 AGCTGCTTATATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1201  
 1254 GATAGGAGAGAGATTTTGGTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1313  
 1202 GATAGGAGAGAGATTTTGGTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1261  
 1314 ATTAACAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1373  
 1262 TTATCTGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1321  
 1374 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1433  
 1322 GGCAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1354  
 1434 AAGATTTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1493  
 1355 TCGATATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1414  
 1494 ACTTTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1553  
 1415 TCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1474  
 1554 AGACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1613  
 1475 TAAAGAGAGAGAGATTAATCACTGATGATGATGATGATGATGATGATGATGATGATGAT 1534  
 1614 GTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1673  
 1535 ATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1594  
 1674 GAAGCCGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1733  
 1595 AATGAGAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1654  
 1734 TTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1779  
 1655 TTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1700

RESULT 15  
 US-09-322-977-3  
 Sequence 3, Application US/09322977  
 Patent No. 6531129  
 GENERAL INFORMATION:  
 APPLICANT: Irls Pecker et al.  
 TITLE OF INVENTION: HEPARANASE SPECIFIC MOLECULAR PROBES  
 TITLE OF INVENTION: AND THEIR USE IN RESEARCH AND MEDICAL  
 NUMBER OF SEQUENCES: 7  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: Mark M. Friedman c/o Anthony Castorina  
 STREET: 2001 Jefferson Davis Highway, Suite 207  
 CITY: Arlington  
 STATE: Virginia  
 COUNTRY: United States of America  
 ZIP: 22202  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
 COMPUTER: Twinhead\* Slimnote-8907X  
 OPERATING SYSTEM: MS DOS version 6.2,  
 SORTWARE: Word for Windows version 3.11  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/322,977  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/922,180

Fri May 14 13:01:53 2004

us-10-088-676-1.rn1

Page 15

FILING DATE: September 2, 1997  
APPLICATION NUMBER: 09/071,739  
FILING DATE: May 1, 1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Fiedemann, Mark W.  
REGISTRATION NUMBER: 33,883  
REFERENCE/DOCKET NUMBER: 910/21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 972-3-562553  
TELEFAX: 972-3-562554  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1721  
TYPE: Nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-09-322-977-3

Query Match 15.7%; Score 278.8; DB 4; Length 1721;  
Best Local Similarity 52.4%; Pred. No. 3.8e-76;  
Matches 684; Conservative 0; Mismatches 592; Indels 30; Gaps 2;

QY 474 AGATTAAGAAAGAGCTGCAAGATGGCCAGCAGCCGATGTTATGCTGCTCCCAAG 533  
DB 425 AGTCAACAGAGATATTTGCAAAATATGATCCATCCCTCTGATGTGAGAGAAATTACG 484  
QY 534 GAGAGAGCAGCTCAGATGATCTGTTCTTCTAAGAGCAATTCCTCAATCTTCA 593  
DB 485 GTTGAATAGGCGCTACAGAGCAATGCTACTCCGAAACATCCAGAAAAAGTTCA 544  
QY 584 TAAATCTATTAACAGCAGAGTCTCTAGCAAACTTTATACCTTGTGCTGCTCTG 653  
DB 545 GAAACAGCCTACTAGAGAGCTGTAGATGCTATACCTTTGCAACCTGCTCAGG 604  
QY 654 ACTCCAGCTGATATTTGCTTAAATGCACTGCTGATTCCTCAATTAATCTCGAAG 713  
DB 605 ACTGAGCTGATCTTTGGCTTAATGCGTTATTAAGAACAGCAATTTGCAAGTGAACG 664  
QY 714 TTCTAGTCCCTGATCTGTTGAAAGTACAGCCGCAAGAAAAAGTCAAACTTTCTGGA 773  
DB 665 TTCTAATGCTCAGTGTCTCTGAGCTACTGCTCTTCCAGGAGTATTAACATTTCTGGA 724  
QY 774 ACTGGATATGAGCCAAATTAATCGAGCAGTATGCGGAGGAGTAAATGCGCAGCA 833  
DB 725 ACTAGGCAATGACCTTAACAGTTCTTAAGAGGCTATATTTCAATGATGGTGCA 784  
QY 834 GTTGGAAAGATTACATCCAGCTGAGAGGCTGTGCAAGCCATCCGATTTATCCAG 893  
DB 785 GTTAGAGAGAGATTAATTCATTCATTAAC---TTCTAAGAAAGTCCACCTTCAAAA 841  
QY 894 AGCCAGCTTAATGAGCCCTTAATATGAGGCGCGAGAGAGATGTCATCGCCCTCTAGA 953  
DB 842 TGCAGAACTCTATGCTGCTGATGTTGTCAGCTTGAAGAAAGAGGCTAAGATGTCGA 901  
QY 954 TGGATTCAAGAGTGGAGAGAGTACAGTACAGTACCTGCAACATTTGTCAT 1013  
DB 902 GAGCTTCTGAAGGCTGGTGGAGAGTATGATTCAGTTACATGGCATCTACTATTT 961  
QY 1014 TGAATGCGGAGTGTGCAAGTGTATGAGTCTCTGAAACTCGCTGTAGACACTCTC 1073  
DB 962 GATAGAGAGAGCTGTACAGAGAGATTTCTTAACCTGATGTATGAGCAATTTTAT 1021  
QY 1074 TGACCAATTAAGAAATTCAGAAAGTGTATTAATACATACCTCCAGAGAAAGATTTG 1133  
DB 1022 TTCACTGTGCAAAAAGTTTCCAGGTGTGAGAGACAGGCTGTGCAAGAGGCTG 1081  
QY 1134 GCTTAAGGTGTGAGCACTCAGCTGAGAGCAAAACATCTATCCGATTCCTATGC 1193  
DB 1082 GTTAGAGAGAAACAGCTCTGATATGAGAGGCGAGCGCCCTTGTATCCGACCTTTGC 1141  
QY 1194 TGCAGATTCTTATGTTGAACATTTAGATGCTGGCAATCAGGGCATGTATGCT 1253

DB 1142 AGCTGCTTTATGCGCTGAGATTAATTGGGCTGTCAAGCCCGAATGGAAATGAAGTGT 1201  
QY 1254 GATACGCACTCATTTTGTGACATGATATACATCACTCTGTGACAGAAATTTAACCC 1313  
DB 1202 GATGAGGCAAGTATCTTTGAGCAGAGAACTACATTTAGTGAATGAATTTGATCC 1261  
QY 1314 ATTACCAATCACTGCTCTCTCTCTCTCAAGGCGCTGATTCGCGCCCAAAAGTTTGGC 1373  
DB 1262 TTACCTGATTTATGCTATCTCTCTCTCAAGAAATTTGGTGGCAACAAAGTGTAT 1321  
QY 1374 TGTGATGTGCTGGGCTCCAGAGAGCAAGGCTGTGCGAGTATCCGAGCAAACT 1433  
DB 1322 GGCAGCGTGAAGTTCAAGAGAGAG-----AAGCT 1354  
QY 1434 AAGATTTATGCTCATGCAACCAACCAACCAACCAACTACGTGCTGCTCATTAAC 1493  
DB 1355 TCGAGTATACCTTCAATGCAACCACTGACAACTCAAGTATTAAGAGAGATTTAAC 1414  
QY 1494 ACTTTTATCATCACTGATTCGATTCAGAAAGAAATCAAGCTGCTGGGACTCTCAG 1553  
DB 1415 TCTGATGCAATTAACCTTCATACGTCACAAAGTATGCGGTTACCTATCTTTTTC 1474  
QY 1554 AGCAAGCTGTTCAACAGTACCTGCTGAGGCTTATGAGGAGAGGCTTAAAGTCAA 1613  
DB 1475 TAAACAGAGTGAATTAATACCTTCTAAGACCTTTGGAGCTCATGATTTCTTCAA 1534  
QY 1614 GTCACTGCACTGATGCGCAGCCCTTATGATGTGAGACAGAGGAGCCCTCCAGATT 1673  
DB 1535 ATCTGTCAACTCAATGATGCTTAACCTTAAGATGTGATGATCAAACTTGCACCTT 1594  
QY 1674 GAAGCCCGCCCTTGGGCGGCGGACATTTGTCATCCCTCCAGTCAACCATGGGCTT 1733  
DB 1595 AATGAGAAAACCTTCGCGCAGAGAGTTCACTGGGCTTGCAGGCTTCTCATATAGTT 1654  
QY 1734 TTATGTGTCAGAAATGTCATGCTTTGGCTGCGGCTTACCATTA 1779  
DB 1655 TTTTGTATGAAGAAATGCAAGTGTGCTGCTTCACTGAAATTA 1700

Search completed: May 13, 2004, 15:40:01  
Job time: 147 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 13, 2004, 09:34:34 ; Search time 6955 Seconds

(without alignments)  
11086.593 Million cell updates/sec

Title: US-10-088-676-1

Perfect score: 1779  
Sequence: 1 atgagggtcgttctgtcctc.....tggcctcgcgcacacgataa 1779

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_ba: \*  
2: gb\_ptg: \*  
3: gb\_in: \*  
4: gb\_on: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vl: \*  
15: em\_ba: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_mu: \*  
20: em\_om: \*  
21: em\_or: \*  
22: em\_ov: \*  
23: em\_pat: \*  
24: em\_ph: \*  
25: em\_pl: \*  
26: em\_ro: \*  
27: em\_sts: \*  
28: em\_un: \*  
29: em\_vl: \*  
30: em\_htg\_hum: \*  
31: em\_htg\_in: \*  
32: em\_htg\_other: \*  
33: em\_htg\_mus: \*  
34: em\_htg\_pln: \*  
35: em\_htg\_rtd: \*  
36: em\_htg\_mam: \*  
37: em\_htg\_vrt: \*  
38: em\_htg\_hum: \*  
39: em\_htg\_mus: \*  
40: em\_htg\_mus: \*  
41: em\_htg\_other: \*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1779	100.0	1779	6	AX100935 Sequence
2	1774.2	99.7	2096	9	AF282887 Homo sapi
3	1774.2	99.7	2353	9	AX1299719 Homo sapi
4	1773.8	99.7	2636	6	AX180830 Sequence
5	1769	99.4	1847	6	AX393920 Sequence
6	1694.2	95.2	1972	9	HS429820
7	1441.4	81.0	1685	6	AX393924
8	1417.8	79.7	1922	9	AF282886
9	1417.8	79.7	2326	6	AX286343
10	1417.8	79.7	2462	6	AX180832
11	1411.4	79.3	2064	9	AX094037
12	1411	79.3	1673	6	AX393922
13	1289.8	72.5	3943	6	AX190449
14	1132.2	63.6	2369	6	AX280370
15	1095.4	61.6	1760	9	AF282885
16	1095.4	61.6	2300	6	AX180834
17	1093.4	61.5	1511	6	AX393926
18	432.4	24.3	468	6	AX259914
19	426.4	24.0	607	6	AX280372
20	375.8	21.1	400	6	AX180851
21	292.8	16.5	1605	5	AY037007
22	290.4	16.3	123482	9	AL445251
23	290.4	16.3	172219	2	AL442004
24	289.2	16.3	1662	4	AF281160
25	285.2	15.0	1736	10	AY077467
26	285.2	15.0	1912	10	AF359507
27	279.4	15.7	1754	10	AF359508
28	278.8	15.7	1721	6	AR080679
29	278.8	15.7	1721	6	AR080680
30	278.8	15.7	1721	6	AR125603
31	278.8	15.7	1721	6	AR125604
32	278.8	15.7	1721	6	AR194189
33	278.8	15.7	1721	6	AR194190
34	278.8	15.7	1721	6	AR221285
35	278.8	15.7	1721	6	AR221286
36	278.8	15.7	1721	6	AR243303
37	278.8	15.7	1721	6	AR243304
38	278.8	15.7	1721	6	AR287435
39	278.8	15.7	1721	6	AR287436
40	278.8	15.7	1721	6	BD074427
41	278.8	15.7	1721	6	BD074428
42	278.8	15.7	1721	6	BD193336
43	278.8	15.7	1721	6	BD193337
44	278.8	15.7	1721	6	BD205238
45	278.8	15.7	1721	6	BD205239

# ALIGNMENTS

RESULT 1  
AX100935  
LOCUS AX100935 1779 bp DNA linear PAT 11-APR-2001  
DEFINITION Sequence 1 from Patent WO0121814.  
ACCESSION AX100935  
VERSION AX100935.1 GI:13619821  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 Duecker, K. and Sirenberg, C.  
Heparanase-2, a member of the heparanase protein family  
Patent: WO 0121814-A 1 29-MAR-2001;  
JOURNAL

FEATURES  
MERCK PATENT GmbH (DE)

Location/Qualifiers  
Source

1. 1779  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
CDS  
1. 1779  
/note="unnamed protein product"

/codon\_start=1  
/protein\_id="CAC36496.1"  
/db\_xref="GI:13619822"  
/translation="MRVLCAPFAPMSSNSRPACAPALVLAHLISLSOAGDR  
RPLPVRNAGLKEKTLILDVSTKRNPRVTNENPLSLDPSI HDGMLEDFSRRLV  
TLARGSPARLRGKRPDLPRONLRNPAKSGRPYIKYKVEDDLYRSDVLDK  
QKCKTIAQHPDVMVLOREKAQHLVLEQPSNTSLILPASLDCIYFADCSG  
LHLIPALNALRNPNSWSSSLSLKTSASRKNVILLDGFKNVAGSTVDAVWG  
SOLKDYIQKSLQPIRIYSRSLYSPNIGRKNVILLDGFKNVAGSTVDAVWG  
HCYIDGRVVKMDPLKTRLDLSDQIRIKOKVNTYTGKIKWEGVTTAGSTNN  
LSDSYAGFLMTLGLANOGIDVIRHSFPHGYNHLVDONFNPDLPLSLYKR  
LIGPKVLAVHAGLQKRPGRVIRDKLIYAHCTHHNHNVDNFIIFILHRSR  
KIKIAGTLADKLHOYLLOPYGSGEIKSKSVQNLQPLVMDDGILPELKRPLRAG  
RIVIPVTMGFVYKXVNALACRYR"

ORIGIN

Query Match 100.0%; Score 1779; DB 6; Length 1779;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1779; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATAGAGGTGCTTTTGCTTCCCTGAAGCCATGCTCCAGCACTCCGCCCCCGCG 60  
DB 1 ATAGAGGTGCTTTTGCTTCCCTGAAGCCATGCTCCAGCACTCCGCCCCCGCG 60  
QY 61 TGCTTACGCGCGGGGCTCTACTTGGCTGTGCTCTCCATCTCCCTTCTCCAG 120  
DB 61 TGCTTACGCGCGGGGCTCTACTTGGCTGTGCTCTCCATCTCCCTTCTCCAG 120  
QY 121 GCTGAGACAGAGACCTTGGCTGTGACAGAGCTGAGTTTGAAGAAAAGACCTG 180  
DB 121 GCTGAGACAGAGACCTTGGCTGTGACAGAGCTGAGTTTGAAGAAAAGACCTG 180  
QY 181 ATTCTACTGATGAGACCAAGAAACCCAGTCAAGACGTCAATGAACTTCTCTCT 240  
DB 181 ATTCTACTGATGAGACCAAGAAACCCAGTCAAGACGTCAATGAACTTCTCTCT 240  
QY 241 CTGCAAGTGGATCGTTCATCATTCATTCATGATGCTGCTGATTCCTTAAGTCG 300  
DB 241 CTGCAAGTGGATCGTTCATCATTCATTCATGATGCTGCTGATTCCTTAAGTCG 300  
QY 301 TTGGTGAACCTTGCGCGGGGACCTTGGCGCGCTTGGCGGGGCAAAAGAAC 360  
DB 301 TTGGTGAACCTTGCGCGGGGACCTTGGCGCGCTTGGCGGGGCAAAAGAAC 360  
QY 361 GACTTCTGCAAGTTCCAGAACTGAGAAACCCGCGCAAAAGCCGCGGGCCGCG 420  
DB 361 GACTTCTGCAAGTTCCAGAACTGAGAAACCCGCGCAAAAGCCGCGGGCCGCG 420  
QY 421 GATTAATGCTCAAAAACCTATGAGATGATTCCTGCAAGATGCTTAAATAA 480  
DB 421 GATTAATGCTCAAAAACCTATGAGATGATTCCTGCAAGATGCTTAAATAA 480  
QY 481 CAGAAAGGCTGCAAGATGCTGCAAGACCTGATGATGCTGCTCAAAAGGAGAG 540  
DB 481 CAGAAAGGCTGCAAGATGCTGCAAGACCTGATGATGCTGCTCAAAAGGAGAG 540  
QY 541 GCGAGCTCAATGATGCTGCTTCTTAAGAGAGCAATTCCTCAATCTTAAGATCTC 600  
DB 541 GCGAGCTCAATGATGCTGCTTCTTAAGAGAGCAATTCCTCAATCTTAAGATCTC 600  
QY 601 ATATTAAAGCAGGCTCTAGACAACTTTAACTTTGATGCTGCTGAGCTCCAG 660  
DB 601 ATATTAAAGCAGGCTCTAGACAACTTTAACTTTGATGCTGCTGAGCTCCAG 660

QY 661 CTGATATTGCTCTAAATGCACTGCGTGTAAATCCCAATACTCTGAAACAGTTCTAGT 720  
DB 661 CTGATATTGCTCTAAATGCACTGCGTGTAAATCCCAATACTCTCTGAAACAGTTCTAGT 720  
QY 721 GCCCGAGTCTGTGAAGTACAGCGGCGAGCAAAAGTAAACATTTTGGGAACTGGGT 780  
DB 721 GCCCGAGTCTGTGAAGTACAGCGGCGAGCAAAAGTAAACATTTTGGGAACTGGGT 780  
QY 781 AATGAGCCAAATTAATCTAGACCATGATGCGCGGCACTTAATGAGCAGCCAGTTGGGA 840  
DB 781 AATGAGCCAAATTAATCTAGACCATGATGCGCGGCACTTAATGAGCAGCCAGTTGGGA 840  
QY 841 AAGGATTAATCCAGCTGAAGAGCGCTTGAAGCCATCCGGAATTTATTCAGAGCCAGC 900  
DB 841 AAGGATTAATCCAGCTGAAGAGCGCTTGAAGCCATCCGGAATTTATTCAGAGCCAGC 900  
QY 901 TTATATGCGCCCTAATATTGGCGCGCGAGGAAAGATGTCATCGCCCTCTAGATGATTC 960  
DB 901 TTATATGCGCCCTAATATTGGCGCGCGAGGAAAGATGTCATCGCCCTCTAGATGATTC 960  
QY 961 ATGAAAGTGGAGAGAGTACAGTATGATGCACTTACCTGCAACATTCCTACATTGATGCG 1020  
DB 961 ATGAAAGTGGAGAGAGTACAGTATGATGCACTTACCTGCAACATTCCTACATTGATGCG 1020  
QY 1021 CGGCTGTCAAGGTGATGCACTTCTGAAACTGCGCTGTAGACACATCTCTGACCCAG 1080  
DB 1021 CGGCTGTCAAGGTGATGCACTTCTGAAACTGCGCTGTAGACACATCTCTGACCCAG 1080  
QY 1081 ATTGAGAAATTCAGAAAGTGTATTAATACATCACTCCAGAAAGAAAGATTGGCTTGA 1140  
DB 1081 ATTGAGAAATTCAGAAAGTGTATTAATACATCACTCCAGAAAGAAAGATTGGCTTGA 1140  
QY 1141 GGTGTGTGACCACTGAGCTGAGAGCAAAACATCTATCCGATTCCTATGCTGACAGGA 1200  
DB 1141 GGTGTGTGACCACTGAGCTGAGAGCAAAACATCTATCCGATTCCTATGCTGACAGGA 1200  
QY 1201 TTCTTATGTTGAACACTTTAGAAATGCTGCGCAATCAAGGCAATGATGCTGATACGG 1260  
DB 1201 TTCTTATGTTGAACACTTTAGAAATGCTGCGCAATCAAGGCAATGATGCTGATACGG 1260  
QY 1261 CACTCATTTTGTGACATGATGATCAATCACTGCTGAGACAGAAATTTTAAACCATTAACA 1320  
DB 1261 CACTCATTTTGTGACATGATGATCAATCACTGCTGAGACAGAAATTTTAAACCATTAACA 1320  
QY 1321 GACTTACTGCTCTCTCTCTCTACCAAGCGCTGATCGGCCCAAGCTTGGCTGTGAT 1380  
DB 1321 GACTTACTGCTCTCTCTCTCTACCAAGCGCTGATCGGCCCAAGCTTGGCTGTGAT 1380  
QY 1381 GTGGCTGGGCTCCAGCGGGAAGCAAGCGCTGGCGGATGATCCGGGACAACTAAGATT 1440  
DB 1381 GTGGCTGGGCTCCAGCGGGAAGCAAGCGCTGGCGGATGATCCGGGACAACTAAGATT 1440  
QY 1441 TATGCTCACTGCAAAACCAACCAACCACTAGTGTGGGTCCATTAACATTTT 1500  
DB 1441 TATGCTCACTGCAAAACCAACCAACCACTAGTGTGGGTCCATTAACATTTT 1500  
QY 1501 ATCATCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
DB 1501 ATCATCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
QY 1561 CTGCTTCAACAGTACCTGCTGACAGCCCTATGAGGAGGAGGCTTAAAGTCCAGTCAAGT 1620  
DB 1561 CTGCTTCAACAGTACCTGCTGACAGCCCTATGAGGAGGAGGCTTAAAGTCCAGTCAAGT 1620  
QY 1621 CAATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680  
DB 1621 CAATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680  
QY 1681 CGCCCGCTTGGGCGGCGGAGCAATGATGATGATGATGATGATGATGATGATGATGATG 1740  
DB 1681 CGCCCGCTTGGGCGGCGGAGCAATGATGATGATGATGATGATGATGATGATGATGATG 1740  
QY 1741 GTCAAGAAATGATGCTTTGGCGCTGCGCTACCAATA 1779

Db 1741 GTCAGAAATGTCATGCTTGGCCCTGCGCTACCGATTA 1779

RESULT 2  
AF282887 2096 bp mRNA linear PRI 14-OCT-2000  
LOCUS AF282887  
DEFINITION Homo sapiens heparanase-like protein HPA2c mRNA, complete cds.  
ACCESSION AF282887  
VERSION AF282887.1 GI:10801198  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 2096)  
McKenzie, E., Tyson, K., Stamps, A., Smith, P., Turner, P., Barry, R., Hircock, M., Patel, S., Barry, E., Stubberfield, C., Terrett, J. and Page, M.  
Cloning and expression profiling of Hpa2, a novel mammalian heparanase family member  
Biochem. Biophys. Res. Commun. 276 (3), 1170-1177 (2000)  
2 (bases 1 to 2096)  
McKenzie, E.A., Tyson, K. and Stamps, A.  
Direct Submission  
Submitted (28-JUN-2000) Biology Group, Oxford Glycosciences, 10 The Quadrant, Abingdon Science Park, Abingdon OX14 3YS, England  
Location/Qualifiers  
source 1..2096  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="10"  
/map="10q23-q24"  
61..1839  
/note="splice variant"  
/codon\_start=1  
/product="heparanase-like protein HPA2c"  
/protein\_id="AA023423.1"  
/db\_xref="GI:10801199"  
/translation="MRVLCAPPEAMPSSNSRPPACALGALYLLALLHLISQAGDR  
RPLVPRAGLKEKTLILLDVSTKPVTVNENPLSLDPSIIHDMLEPLSEKRLV  
TLARGLSPALFPGKRTDFQFONLRAPKSRGPGEDILKLYEDIVASDVALDK  
OKGKTIADHPVMELOREKXAOHMLVLKXOFNSNTSLILTRSDIKYFNADCG  
LHLIPALAIRPNPNSSSSALSILKYSKRYNISMEIGNEPNRYRTMHGAVNG  
SOLGDDYOLKSLIOPRIYRSASISGPNIGRPKNYIALLDGEMKTAGSYDATTG  
HCTIDGRVYKWDPLKRLIDPLSLDQIRKIQKVVNTTPGKRLHSGVTTISAGTIN  
LSDSYAAGFLWNLGMLANGIDIVIRHSFDFGYNHLVDQNFPLHGVTTLSLRYR  
LIGPRLVAVHAGLQKRPGRVIRDKLRIYAHCTNHNHNVGASITLPIINDHRSR  
KKIKLAGTLRDLVHOYLLOPYGQEGELKSKSVQANGQPLVWVDGT.PELKPRPLRAG  
RLVLPPTMGFFVKNVNALACRRR"

ORIGIN

Query Match 99.7%; Score 1774.2; DB 9; Length 2096;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1776; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGAGGAGTCTTGGCTTCCCTGAAGCAATGCCCTCAGACAACCTCCGCGCGCGCG 60  
Db 61 ATGAGGAGTCTTGGCTTCCCTGAAGCAATGCCCTCAGACAACCTCCGCGCGCGCG 120  
QY 61 TGCCTAGCCCGCGGAGCTCTCTACTGCTCTGTGCTCAATCTCTCCCTTCTCCAG 120  
Db 121 TGCCTAGCCCGCGGAGCTCTCTACTGCTCTGTGCTCAATCTCTCCCTTCTCCAG 180  
QY 121 TGCCTAGCCCGCGGAGCTCTCTACTGCTCTGTGCTCAATCTCTCCCTTCTCCAG 180  
Db 121 GCTGAGACAGAGACCCCTGCTCTAGACAGAGCTGAGGTTGAAGAAAAGACCTTG 240  
QY 181 GCTGAGACAGAGACCCCTGCTCTAGACAGAGCTGAGGTTGAAGAAAAGACCTTG 240  
Db 181 ATTCTACTGATGTAGACCAAGAACCCAGTCAGAGACGTCAATGAGAACTTCTCTCT 240

Db 241 ATTCTACTGATGTAGACCAAGAACCCAGTCAGAGACGTCAATGAGAACTTCTCTCT 300  
QY 241 CTGACGCTGGATCCGTCATCATTTGATGAGTGGCTGATTTCTTAAGCTCAAGCGC 300  
Db 301 CTGACGCTGGATCCGTCATCATTTGATGAGTGGCTGATTTCTTAAGCTCAAGCGC 360  
QY 301 TTGCTGACCTTGAGCCCGGAGACTTTTGGCCGCTTTCTGCGCTCGGGGCAAGAGACC 360  
Db 361 TTGCTGACCTTGAGCCCGGAGACTTTTGGCCGCTTTCTGCGCTCGGGGCAAGAGACC 420  
QY 361 GACTTCCTGAGTCCAGAACCTGAGAGAACCCGCGGAAAAGCGCGGGGCGCGCGCG 420  
Db 421 GACTTCCTGAGTCCAGAACCTGAGAGAACCCGCGGAAAAGCGCGGGGCGCGCGCG 480  
QY 421 GATTACTATCTCAAAAAGTATGAGATGACATTTGTCGAAGTATGTTGCTTGAATAA 480  
Db 481 GATTACTATCTCAAAAAGTATGAGATGACATTTGTCGAAGTATGTTGCTTGAATAA 540  
QY 481 CAGAAAGGCTGCAAGATGCGCCGACACCTGATGTTATGCTGCTCCAAAGGGAAG 540  
Db 541 CAGAAAGGCTGCAAGATGCGCCGACACCTGATGTTATGCTGAGCTCCAAAGGGAAG 600  
QY 541 GCAGCTCAGATGATGCTGCTTCTTCTTAAAGAGCAATTCATTAAGTATCTC 600  
Db 601 GCAGCTCAGATGATGCTGCTTCTTCTTAAAGAGCAATTCATTAAGTATCTC 660  
QY 601 ATATTACAGCCAGGCTCTAGACAACTTTATTAATCTTGTGATGCTCTGAGCTCAC 660  
Db 661 ATATTACAGCCAGGCTCTAGACAACTTTATTAATCTTGTGATGCTCTGAGCTCAC 720  
QY 661 CTGATATTTGCTTAATGACATGCGCTGCTTAATCCATTAATCTTGAAGTCTAGT 720  
Db 721 CTGATATTTGCTTAATGACATGCGCTGCTTAATCCATTAATCTTGAAGTCTAGT 780  
QY 721 GCGCTGAGTCTGTTGAAGTACAGCGCCGACCAAAAAGTATCAACTTTCTTGGAACTGGT 780  
Db 781 GCGCTGAGTCTGTTGAAGTACAGCGCCGACCAAAAAGTATCAACTTTCTTGGAACTGGT 840  
QY 781 AATGAGCAAAATACATATGAGACCAATGAGCGCGGAGTAAATGAGCGCAAGTGGGA 840  
Db 841 AATGAGCAAAATACATATGAGACCAATGAGCGCGGAGTAAATGAGCGCAAGTGGGA 900  
QY 841 AAGGATTAATCAATCAAGTGAAGAGCTGTTGACAGCCCATCCGATTTATTCAGAGCGAG 960  
Db 901 AAGGATTAATCAATCAAGTGAAGAGCTGTTGACAGCCCATCCGATTTATTCAGAGCGAG 960  
QY 901 TTATATGGCCCTAATATTTGGCGCGCGAGAGAAAGTATGATGCGCTCTAGATGATTC 960  
Db 961 TTATATGGCCCTAATATTTGGCGCGCGAGAGAAAGTATGATGCGCTCTAGATGATTC 1020  
QY 961 ATGAGAGTGGCAGAGAGTACAGTATGAGTACCTGTCACAACTTGTCTACATTTGATGCG 1020  
Db 1021 ATGAGAGTGGCAGAGAGTACAGTATGAGTACCTGTCACAACTTGTCTACATTTGATGCG 1080  
QY 1021 CCGGCTGCTCAAGTATGAGTATCTCTGAAAATCTGCGCTTGTAGACACATCTCTGACCG 1080  
Db 1081 CCGGCTGCTCAAGTATGAGTATCTCTGAAAATCTGCGCTTGTAGACACATCTCTGACCG 1140  
QY 1081 ATTAGAGAAATTCAGAAATGCTTTAATACATACATCTCCAGAAAAGAAATTTGGCTTGA 1140  
Db 1141 ATTAGAGAAATTCAGAAATGCTTTAATACATACATCTCCAGAAAAGAAATTTGGCTTGA 1200  
QY 1141 GGTGTGATGACCACTCACTGAGGAGCACAAAACATCTATCCGATTTCTATGCTGACGA 1200  
Db 1201 GGTGTGATGACCACTCACTGAGGAGCACAAAACATCTATCCGATTTCTATGCTGACGA 1260  
QY 1201 TTCTTATGCTGAACATTTAGAAATGCTGGCAATCAGGGCATGATGATGATGATGATGAT 1260  
Db 1261 TTCTTATGCTGAACATTTAGAAATGCTGGCAATCAGGGCATGATGATGATGATGATGAT 1320  
QY 1261 CACTCATTTTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
Db 1321 CACTCATTTTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380





QY	961	ATGAAGGTGGAGGAAAGTACAGTAGATCAGTTACCTGGCAACATTTGCTACATTGATGGC	1028
Db	1021	ATGAAGGTGGAGGAAAGTACAGTAGATCAGTTACCTGGCAACATTTGCTACATTGATGGC	1088
QY	1021	CGGGTGTGCMAAGTATGAGACTTCTGTAAAACCTGGCTTTAGACACACTCTGTACCAAG	1088
Db	1081	CGGGTGTGCMAAGTATGAGACTTCTGTAAAACCTGGCTTTAGACACACTCTGTACCAAG	1140
QY	1081	ATTAGGAATAATCAGAAAGTGGTTATATACATCACTCCAGGAAGAATAATTTGGCTTGA	1140
Db	1141	ATTAGGAATAATCAGAAAGTGGTTATATACATCACTCCAGGAAGAATAATTTGGCTTGA	1200
QY	1141	GGTGTGTGACCACTCAGCTGAGGACCAAAACAATCTATCCGATTTCTTATGCTTCAGGA	1200
Db	1201	GGTGTGTGACCACTCAGCTGAGGACCAAAACAATCTATCCGATTTCTTATGCTTCAGGA	1260
QY	1201	TTCTATATGGTTGAACACTTTTAGAATGCTGGCCAAATCAAGGGCATTTAGTGTGATACGG	1260
Db	1261	TTCTATATGGTTGAACACTTTTAGAATGCTGGCCAAATCAAGGGCATTTAGTGTGATACGG	1320
QY	1261	CACCTCAATTTTGTGACCAATGAGATACATCACTCGTAGACACAGAAATTTTAAACCAATTA	1320
Db	1321	CACCTCAATTTTGTGACCAATGAGATACATCACTCGTAGACACAGAAATTTTAAACCAATTA	1380
QY	1321	GACATCACTGGCTCTCTCTCCTCTACAAAGGACCTGATCCGGCCCAAAAGCTTGTGGCTGTGAT	1380
Db	1381	GACATCACTGGCTCTCTCTCCTCTACAAAGGACCTGATCCGGCCCAAAAGCTTGTGGCTGTGAT	1440
QY	1381	GTGGCTGGGCTCCAGCGGAAAGCCACAGGCTTGGCCGAGTATCCGGGACAAACTTAAGATT	1440
Db	1441	GTGGCTGGGCTCCAGCGGAAAGCCACAGGCTTGGCCGAGTATCCGGGACAAACTTAAGATT	1500
QY	1441	TATGTCTCACTGACAAACAACAACAACAACAATCTATCGTGGCTTCATTATACCTTTT	1500
Db	1501	TATGTCTCACTGACAAACAACAACAACAACAATCTATCGTGGCTTCATTATACCTTTT	1560
QY	1501	ATCATCAACTTGTGCATGCATCAAGAAGAAATCAAGCTGGGTGGGACCTCTAGAGACAAG	1560
Db	1561	ATCATCAACTTGTGCATGCATCAAGAAGAAATCAAGCTGGGTGGGACCTCTAGAGACAAG	1620
QY	1561	CTGGTTCAACCAATACCTGCTGTGCAGCCCTATGGGCGAGAGGCTTAAATGCCAGTCAATG	1620
Db	1621	CTGGTTCAACCAATACCTGCTGTGCAGCCCTATGGGCGAGAGGCTTAAATGCCAGTCAATG	1680
QY	1621	CAACTGAATYGGCCAGACCCCTTAGTATYAGTGAAGCAAGGAGACCCCTCCAGAAATTTGAAGCC	1680
Db	1681	CAACTGAATYGGCCAGACCCCTTAGTATYAGTGAAGCAAGGAGACCCCTCCAGAAATTTGAAGCC	1740
QY	1681	CGCCCCCTTTCGGGCGCGGCGGACATTTGTCATCCTCCAGTCAACATGGGCTTTTATGTG	1740
Db	1741	CGCCCCCTTTCGGGCGCGGCGGACATTTGTCATCCTCCAGTCAACATGGGCTTTTATGTG	1800
QY	1741	GTCAGGAATGTCAATGCTTTGGCCCTGGCCGCAACCGATA	1779
Db	1801	GTCAGGAATGTCAATGCTTTGGCCCTGGCCGCAACCGATA	1839
RESULT 4			
AX180830			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
REFERENCE			
1			
McKenzie, E. A., Stamps, A. C., Terrett, J. A., and Tyson, K. L.			
Homologues of human heparanase and splice variants thereof			
Patent: WO 0146392-A 1 28-JUN-2001;			
JOURNAL			

FEATURES	Source	Location/Qualifiers
Oxford Glycosciences (UK) Limited (GB)		
1..2636		/organism="Homo sapiens"
		/mol_type="unassigned DNA"
		/db_xref="taxon:9606"
		601..2379
		/note="unamed protein product"
		/codon_start=1
		/protein_id="CAC50008.1"
		/db_xref="GI:15132667"
		/db_xref="REMBL: CAC50008"
		/translation="MRYLCAPFEMAPESNSRPPACLPALVYALILHLISLSQAGDPR RPLPVDPAAGLKEKTLILLDVSTPNPRTVNEMLSLQDPSIITHDMLPESKRLVR TLAAGSPAFPLFEGGKSTDFLOFNILKPNVSKRGGPDYKIKNEDIVRSDDYDCKG OKGOKIAHPDVMLEQREKAQOMHVLKQEPNYSNLIITRPSIDLPVNDPDSG LHLIFALNALRRPNNSWSSSALSLKYSASKRYNLSWELGNEPNRYTRMGAAVNG SQUKDTQLKSLQPLQIRITRSALVYEPNIGRPKNVYALLDGPMKTAGSTVDVTVQ HCYIDKSVKVMDFLKTRLIDTLSDQIRKI QKVNTYTPGKMLLEGVVTSAGINN LSBSYAGPLMLNTLGMLANQIGIDVVRHSFPHGVNHLVDQNEMLPDDVLSLYKR LIGSKYLAIVAVAGLQKRPGRGVIRDLRIYAHCTNNHNNYVGSITLFINHRSR KKIILAGTLDKLVHQLQLOPYGQESKSVQMGQPLVWVDGTLPELKPRLRAG RTIVIPVTMGFVVKVNMLACRYR"
ORIGIN		
Query Match	99.7%	Score 1773.8; DB 6; Length 2636;
Best Local Similarity	99.8%	Pred. No. 0;
Matches 1775; Conservative	0; Mismatches	4; Indels 0; Gaps 0;
1	ATGAGGGTGCTTTTGCCCTTCCTCGAAGCCATGCGCTCCAGCACTCCGCGCGCGCG	60
601	ATGAGGGTGCTTTTGCCCTTCCTCGAAGCCATGCGCTCCAGCACTCCGCGCGCGCG	660
61	TGCGTAGCCCGGGGGGCTCTCTACTTGCTGTGTGCTGCATCTCCCTTCTCCAG	120
661	TGCGTAGCCCGGGGGGCTCTCTACTTGCTGTGTGCTGCATCTCCCTTCTCCAG	720
121	GCTGGAACAGAGAACCTTTCGCTGTGACAGACACTGACGTTTGAAAGAAACCTTG	180
721	GCTGGAACAGAGAACCTTTCGCTGTGACAGACACTGACGTTTGAAAGAAACCTTG	780
181	ATTCTACTGATGTGAGACACCAAGACCCAGTCAAGACAGTCATGTGAATCTCTCT	240
781	ATTCTACTGATGTGAGACACCAAGACCCAGTCAAGACAGTCATGTGAATCTCTCT	840
241	CTGCAGCTGATCCGTCATCATTCATGATGCTGCTGATTCCTTAAGCTCCAAAGCG	300
841	CTGCAGCTGATCCGTCATCATTCATGATGCTGCTGATTCCTTAAGCTCCAAAGCG	900
301	TTGGTGACCCCGGCGCGGGGACTTTCGCGCGCGCTTTCGCGCTTCGCGGGGCAAAAGGACC	360
901	TTGGTGACCCCGGCGCGGGGACTTTCGCGCGCGCTTTCGCGCTTCGCGGGGCAAAAGGACC	960
361	GACTTCTGCACTTCAGAACCTTGAGAGAACCCCGCGGAAAAGCCGCGGGGCGCGCGCG	420
961	GACTTCTGCACTTCAGAACCTTGAGAGAACCCCGCGGAAAAGCCGCGGGGCGCGCGCG	1020
421	GATTACATATCTCAAAAATCTATGAGATGACATTGTTGGAAGTGAATTTGGCTTGATPAA	480
1021	GATTACATATCTCAAAAATCTATGAGATGACATTGTTGGAAGTGAATTTGGCTTGATPAA	1080
481	CAGAAAGGCTGCAAGATTGGCCAGACCCCTGATTTATGCTGTGCTCCAAAGGAGAGAG	540
1081	CAGAAAGGCTGCAAGATTGGCCAGACCCCTGATTTATGCTGTGCTCCAAAGGAGAGAG	1140
541	GCACTCAAGATGCATCTGGTTCCTCTAAGAGCAATTCCTCAATACTTACAGTAATCTC	600
1141	GCACTCAAGATGCATCTGGTTCCTCTAAGAGCAATTCCTCAATACTTACAGTAATCTC	1200
601	ATATTAACAGCAGAGTCTAGACAAACTTTATACCTTGGTGAATGCTCTGAGACTCCAC	660
1201	ATATTAACAGCAGAGTCTAGACAAACTTTATACCTTGGTGAATGCTCTGAGACTCCAC	1260





QY 361 GACTTCCTGAGTTCAGAGAACTGAGAAACCCGCGAGAAAAGCCGCGGCGCCGCGGCGG 420  
 Db 415 GACTTCCTGAGTTCAGAGAACTGAGAAACCCGCGAGAAAAGCCGCGGCGCCGCGGCGG 474  
 QY 421 GATTACTATCTCAAAAACCTATGAGATGACATTGTTGAGAGTGAATGTTGCTTATGATTA 480  
 Db 475 GATTACTATCTCAAAAACCTATGAGATGACATTGTTGAGAGTGAATGTTGCTTATGATTA 534  
 QY 481 CAGAAAGGCTGCAAGATTTGCCAGACCTGATGTTATGCTGATGCTGCAAGGAGAGAG 540  
 Db 535 CAGAAAGGCTGCAAGATTTGCCAGACCTGATGTTATGCTGATGCTGCAAGGAGAGAG 594  
 QY 541 GCAGCTCAGATGATCTGAGTCTTCTTAAGAGACATTTCTCAATTAACATTAACATTA 600  
 Db 595 GCAGCTCAGATGATCTGAGTCTTCTTAAGAGACATTTCTCAATTAACATTAACATTA 654  
 QY 601 ATATTAAAGCCAGGCTCTTGAACAATTTTAACTTTGCTGATGCTGCTGAGCTGAC 660  
 Db 655 ATATTAAAGCCAGGCTCTTGAACAATTTTAACTTTGCTGATGCTGCTGAGCTGAC 714  
 QY 661 CTGATATTGCTCTTAATGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 720  
 Db 715 CTGATATTGCTCTTAATGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 774  
 QY 721 GCCCTGAGTCTGTTGAAGTACAGCGCCAGCAAAAAGTACAACTTTCTTGGGAACTGGGT 780  
 Db 775 GCCCTGAGTCTGTTGAAGTACAGCGCCAGCAAAAAGTACAACTTTCTTGGGAACTGGGT 834  
 QY 781 AATGAGCAAAATTAATATGAGCAATGATGAGCGCGGAGCAAAATGAGCAAGTGGGA 840  
 Db 835 AATGAGCAAAATTAATATGAGCAATGATGAGCGCGGAGCAAAATGAGCAAGTGGGA 894  
 QY 841 AAGATTAATATGAGCAATGATGAGCGCGGAGCAAAATGAGCAAGTGGGA 900  
 Db 895 AAGATTAATATGAGCAATGATGAGCGCGGAGCAAAATGAGCAAGTGGGA 954  
 QY 901 TTAATGAGCAATGATGAGCGCGGAGCAAAATGAGCAAGTGGGA 960  
 Db 955 TTAATGAGCAATGATGAGCGCGGAGCAAAATGAGCAAGTGGGA 1014  
 QY 961 ATGAGAGTGGCAGAGAGTACAGTATGATGATGATGATGATGATGATGATGATGATG 1020  
 Db 1015 ATGAGAGTGGCAGAGAGTACAGTATGATGATGATGATGATGATGATGATGATGATG 1074  
 QY 1021 CGGGTGGTCAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080  
 Db 1075 CGGGTGGTCAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1134  
 QY 1081 ATTAGAGAAATTCAGAAAGTGTATATATATATATATATATATATATATATATATAT 1140  
 Db 1135 ATTAGAGAAATTCAGAAAGTGTATATATATATATATATATATATATATATATATAT 1194  
 QY 1141 GGTGGTGGTCAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200  
 Db 1195 GGTGGTGGTCAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1254  
 QY 1201 TTTCTATGTTGAACCTTTAGAAAGTGGCAATCAAGGGGCAATGATGATGATGATGATG 1260  
 Db 1255 TTTCTATGTTGAACCTTTAGAAAGTGGCAATCAAGGGGCAATGATGATGATGATGATG 1314  
 QY 1261 CACTCATTTTGGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320  
 Db 1315 CACTCATTTTGGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1374  
 QY 1321 GACTATGAGTCT 1380  
 Db 1375 GACTATGAGTCT 1434  
 QY 1381 GTGGCTGGGCTTCAGCGGAGCAAGCGGCTGGCCGAGTATCGGAGCAAAATTAAGATT 1440  
 Db 1435 GTGGCTGGGCTTCAGCGGAGCAAGCGGCTGGCCGAGTATCGGAGCAAAATTAAGATT 1494  
 QY 1441 TATGCTCACTGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCACTTTT 1500

Db 1495 TATGCTCACTGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCACTTTT 1554  
 QY 1501 ATCATCAACTTGCATGATCAAGAAAGAAATCAAGCTGGCTGGGACTCTGAGAGCAAG 1560  
 Db 1555 ATCATCAACTTGCATGATCAAGAAAGAAATCAAGCTGGCTGGGACTCTGAGAGCAAG 1614  
 QY 1561 CTGGTTCACAGTACCTGCTGAGCGCTTATGAGTGAACGAGAGGCTTAAGTCCAAATGCTG 1620  
 Db 1615 CTGGTTCACAGTACCTGCTGAGCGCTTATGAGTGAACGAGAGGCTTAAGTCCAAATGCTG 1674  
 QY 1621 CAACTGAATGGCAGCGCTTATGAGTGAACGAGAGGCTTAAGTCCAAATGCTGAG 1680  
 Db 1675 CAACTGAATGGCAGCGCTTATGAGTGAACGAGAGGCTTAAGTCCAAATGCTGAG 1734  
 QY 1681 CGCCCTCTGCGGCGCGGAGCAATGATGATGATGATGATGATGATGATGATGATGATG 1740  
 Db 1735 CGCCCTCTGCGGCGCGGAGCAATGATGATGATGATGATGATGATGATGATGATGATG 1794  
 QY 1741 GTCAAGAAATGCAATGCTTTGGCTGCGCGCTACCGATTA 1779  
 Db 1795 GTCAAGAAATGCAATGCTTTGGCTGCGCGCTACCGATTA 1833  
 RESULT 6  
 HSA299720 1872 bp mRNA linear PRI 04-JUN-2002  
 LOCUS  
 DEFINITION Homo sapiens mRNA for heparanase 3.  
 ACCESSION AJ299720  
 VERSION AJ299720.1 GI:18073438  
 KEYWORDS heparanase 3.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 1 Legoux P., Legoux R., O'Brien D. and Salome M.  
 Unpublished  
 2 (bases 1 to 1872)  
 JOURNAL  
 AUTHORS Pessegue Safontas, B.J.O.P.S.  
 TITLE Direct Submission  
 JOURNAL Submitted (18-SEP-2000) Pessegue Safontas B.J.O.P.S., Bioinformatics, Sanofi-Synthelabo, Sanofi Recherche - immopole - Voie 1 - BP 137, 31676 Labège CEDEX, FRANCE  
 FEATURES  
 source  
 1..1872  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="prostate"  
 1..1647  
 /codon\_start=1  
 /product="heparanase 3"  
 /protein\_id="ACB2492.1"  
 /db\_xref="GI:18073439"  
 /translation="MRVLCAPFPMILSSNSRPACIAPALYIALILHLSLSQADR  
 RLPLVPAAGLKEKTLILDLVSKNVRVVENFSLQDPSLIHGWIDFLSSKLV  
 TLARGSPAFRLFGKRTDPLQFONLRNPKSGGSPDYLYLNYEDDIVRSDVALDK  
 QKCKIAHDPDWLELQREKAAOMHLVLEKQPSNYSNLILTARSLDKLYNSADCSG  
 IHLIFPINALRPNPNKSNSSALSILKYSASKVNI SWELNIEPNVYRSGRAVNG  
 SGLKRYIQLKSLIQPIRYSPASLYSPNIGPRKRVIALLGPMVAVASTYDAATMO  
 HCYIDRVRKMDFLTRLDLTDSDIRIKI QKVNTYITFGKIMLEGVYTTSGATNN  
 LQSVYAGFLMNTLGLMNLQGDIVIRISFDHGNHLVDQNFNLPDYWLSLTKR  
 LIQKVLAVHVAQIQRPGRVIRDKAIYVACCTNHNHNVYRSGITLFIINLHRSR  
 KIKIAGTLRDKLVHQLYLLQPYQEBLKSRTQCYGII"  
 ORIGIN  
 Query Match 95.2%; Score 1694.2; DB 9; Length 1872;  
 Best Local Similarity 96.1%; Pred. No. 0;  
 Matches 1776; Conservative 0; Mismatches 3; Indels 70; Gaps 1;  
 QY 1 ATGAGAGTGGCTTTGCTTCTCTGAAAGCCATCCTTCAGCAACTCCGCGCCCGCGG 60

Db 1 ATGAGGATGCTTTGCTTCCCTGAGGCAATGCTCCAGCAACTCCGCCCCCGG 60  
 Qy 61 TGGCTAGCCCCGGGGCTCTCTAATTGGCTCTGCTGCTCACTTCCCTTTCTCCAG 120  
 Db 61 TGGCTAGCCCCGGGGCTCTCTAATTGGCTCTGCTGCTCACTTCCCTTTCTCCAG 120  
 Qy 121 GCTGAGACAGAGAACCTTGGCTGTAGACAGAGCTGACAGGTTTGAAGAAAGACCTG 180  
 Db 121 GCTGAGACAGAGAACCTTGGCTGTAGACAGAGCTGACAGGTTTGAAGAAAGACCTG 180  
 Qy 181 ATTCTATCTGATGTAGACCAAGAACCCAGTACAGACAGTCAATGAGAACTTCTCTCT 240  
 Db 181 ATTCTATCTGATGTAGACCAAGAACCCAGTACAGACAGTCAATGAGAACTTCTCTCT 240  
 Qy 241 CTGAGAGTGGATCCGTTCATCATTCATGATGGCTGGCTGATTTCTTAAGCTCCAGCGG 300  
 Db 241 CTGAGAGTGGATCCGTTCATCATTCATGATGGCTGGCTGATTTCTTAAGCTCCAGCGG 300  
 Qy 301 TTGGTGAACCTGAGCCCGGGGACTTTCCGCGCTTTCTGCGCTTCGGGGGCAAAAGGACC 360  
 Db 301 TTGGTGAACCTGAGCCCGGGGACTTTCCGCGCTTTCTGCGCTTCGGGGGCAAAAGGACC 360  
 Qy 361 GACTTCTGCACTTCCAGAACTGAGAGAACCCGGGCGAAAGCCGGGGGGCCGGCGCG 420  
 Db 361 GACTTCTGCACTTCCAGAACTGAGAGAACCCGGGCGAAAGCCGGGGGGCCGGCGCG 420  
 Qy 421 GATTACTATCTCAAAAATATGAGAGTACATTTGTGAGAGTGTGCTTGAATAA 480  
 Db 421 GATTACTATCTCAAAAATATGAGAGTACATTTGTGAGAGTGTGCTTGAATAA 480  
 Qy 481 CAGAAAGGCTGCAAGATTGCCAGACACCTGATGTATGCTGAGGCTCCAAAGGAGAA 540  
 Db 481 CAGAAAGGCTGCAAGATTGCCAGACACCTGATGTATGCTGAGGCTCCAAAGGAGAA 540  
 Qy 541 GCAGCTGAGATGATCTGCTCTCTTAAGAGCAATTCCTCAATCTCACTACATCTC 600  
 Db 541 GCAGCTGAGATGATCTGCTCTCTTAAGAGCAATTCCTCAATCTCACTACATCTC 600  
 Qy 601 ATATTACAGCCAGGCTCTCTGACAACTTTATTAATCTTGTGATGTCTGACACTCC 660  
 Db 601 ATATTACAGCCAGGCTCTCTGACAACTTTATTAATCTTGTGATGTCTGACACTCC 660  
 Qy 661 CTGATATTGCTCTAAATGCACTGCTGCTGATGCTGATGCTGATGCTGATGCTGAT 720  
 Db 661 CTGATATTGCTCTAAATGCACTGCTGCTGATGCTGATGCTGATGCTGATGCTGAT 720  
 Qy 721 GCCCTGAGCTGCTGAGAGTACAGGCGCAGCAAAAAGTACAACTTTCTGGGAACTGG 780  
 Db 721 GCCCTGAGCTGCTGAGAGTACAGGCGCAGCAAAAAGTACAACTTTCTGGGAACTGG 780  
 Qy 781 AATGAGCAATTAATCTATCCAGCACTGATGCGGGGAGTAAATGAGAGAGCTTGGGA 840  
 Db 781 AATGAGCAATTAATCTATCCAGCACTGATGCGGGGAGTAAATGAGAGAGCTTGGGA 840  
 Qy 841 AAGGATTAATCACTCACTGAGAGAGCTGCTGAGAGAGCTGAGATTTATTCAGAGCCAG 900  
 Db 841 AAGGATTAATCACTCACTGAGAGAGCTGCTGAGAGAGCTGAGATTTATTCAGAGCCAG 900  
 Qy 901 TTATATGCTGCTTAATTTGGGGGCGGAGAGAAATGCAATGCTGCTGCTGAGATGAT 960  
 Db 901 TTATATGCTGCTTAATTTGGGGGCGGAGAGAAATGCAATGCTGCTGCTGAGATGAT 960  
 Qy 961 ATGAGGTGAGAGAGTACAGTATGAGTACAGTATGAGTATGAGTATGAGTATGAGT 1020  
 Db 961 ATGAGGTGAGAGAGTACAGTATGAGTACAGTATGAGTATGAGTATGAGTATGAGT 1020  
 Qy 1021 CCGGCTGCTGAGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 1080  
 Db 1021 CCGGCTGCTGAGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 1080  
 Qy 1081 ATTAGAAATTCAGAAAGTGTATATCATCACTCCAGAGAAAGATTTGGCTTGA 1140

Db 1081 ATTAGAAATTCAGAAAGTGTATATCATCACTCCAGAGAAAGATTTGGCTTGA 1140  
 Qy 1141 GGTGTGGAGACCACTCAAGCTGAGAGGCAAAAGATCATCCATCTTCTATGCTGAGGA 1200  
 Db 1141 GGTGTGGAGACCACTCAAGCTGAGAGGCAAAAGATCATCCATCTTCTATGCTGAGGA 1200  
 Qy 1201 TTCTTATGTTGAAACATTTAGAGATGCTGGCAATGAGGATGATGCTGATACGG 1260  
 Db 1201 TTCTTATGTTGAAACATTTAGAGATGCTGGCAATGAGGATGATGCTGATACGG 1260  
 Qy 1261 CACTCATTTTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
 Db 1261 CACTCATTTTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
 Qy 1321 GACTATGCTCTCTCTCTCTCAAGAGGCTGATGAGGCTCCCAAAATCTTGGCTGAT 1380  
 Db 1321 GACTATGCTCTCTCTCTCTCAAGAGGCTGATGAGGCTCCCAAAATCTTGGCTGAT 1380  
 Qy 1381 GTGGCTGGGCTTCCAGCGGAGACCAAGGCTGGCGGAGTCCGGAGCAATAGAGATT 1440  
 Db 1381 GTGGCTGGGCTTCCAGCGGAGACCAAGGCTGGCGGAGTCCGGAGCAATAGAGATT 1440  
 Qy 1441 TATGCTCACTGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACTTT 1500  
 Db 1441 TATGCTCACTGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACTTT 1500  
 Qy 1501 ATCATCACTTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
 Db 1501 ATCATCACTTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
 Qy 1561 CTGTTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620  
 Db 1561 CTGTTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620  
 Qy 1621 AGATGTCATATCTGAGATCATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
 Db 1621 AGATGTCATATCTGAGATCATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
 Qy 1681 ---GTCAAGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
 Db 1681 ---GTCAAGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
 Qy 1741 ATTGAAGCCCGCCCTTCCGGGCGGCGGAGATGATGATGATGATGATGATGATGAT 1800  
 Db 1741 ATTGAAGCCCGCCCTTCCGGGCGGCGGAGATGATGATGATGATGATGATGATGAT 1800  
 Qy 1731 CTTTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1779  
 Db 1801 CTTTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1849

RESULT 7  
 AX393924 1685 bp DNA linear PAT 23-MAR-2002  
 LOCUS AX393924  
 DEFINITION Sequence 5 from Patent WO0204645.  
 ACCESSION AX393924  
 VERSION AX393924.1 GI:19701884  
 KEYWORDS  
 SOURCE  
 ORGANISM

Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE  
 AUTHORS David, G., and Duerr, J.  
 TITLE A second human heparanase, and splice variants thereof, with a  
 predominant expression in skeletal muscle, heart and pancreas  
 Patent: WO 0204645-A 5 17-JAN-2002;  
 JOURNAL Vlaams Interuniversitair Instituut voor Biotechnologie vzw. (BE)

FEATURES  
 source  
 1..1685  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

CDS

85..1671  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD29061.1"  
 /db\_xref="gi:19701885"  
 /db\_xref="EMBL:CAD29061"  
 /translation="MPSNSRPACTIAGALYLAHLHLSLSOAGDRRLPVDRAG  
 LKEXLILIDVSTKNPVTVNENFLSLDPSLIHDEGLDLSKRLVTLARGLSPAF  
 LAFGRKADFLPOPLRNPARKSRRGSPDYLYKYEDPARSLDKYLPADCSGLIFA  
 LNALRNPNSNWSALSLKISASKKTNISWELGNPNPNRTMGRANVSGLDGD  
 YIOLKSLQPIRISYRSLYGNIGRPKRVIALDGFVKVASTVDAVTVQCTYIG  
 RVKTMDFKRLDLDSQIDQIKIOKVNTYTPGKXIMLEAVTVTSAGTNLSDSA  
 AGFLMNTLGMLANQIDIVILHSFDFHGNHMLVDONENPLPDYMLSLYRLLGPVY  
 LAVHAGLORXKRPGRVIRADKLIRYACNNHNNHYVSGSLTFLINLSRKKIKLA  
 GTLRKLVHOLYLPYGOGLKSKSVQNLQPLVWVDGTLPELKPRLRAGRTLVIP  
 PVTMGFPVKNVNALACRYR"

ORIGIN

Query Match 81.0%; Score 1441.4; DB 6; Length 1685;

Best Local Similarity 90.7%; Pred. No. 0;

Matches 1614; Conservative 1; Mismatches 2; Indels 162; Gaps 1;

QY 1 ATGAGGGTCTTTTGTGCTTCCCTGAAGCCATGCCCTCAGAACTCCGCGCCCGCGCG 60  
 DB 55 ATGAGGGTCTTTTGTGCTTCCCTGAAGCCATGCCCTCAGAACTCCGCGCCCGCGCG 114  
 QY 61 TGGCTAGCCCGGGGGCTCTACTTGTGCTGCTCATCTCTCCCTTCTCCGAG 120  
 DB 115 TGGCTAGCCCGGGGGCTCTACTTGTGCTGCTCATCTCTCCCTTCTCCGAG 174  
 QY 121 GCTGAGACAGAGACCTTGGCTGTAGACAGAGCTGCAAGTTTGAAGAAAAGCCTG 180  
 DB 175 GCTGAGACAGAGACCTTGGCTGTAGACAGAGCTGCAAGTTTGAAGAAAAGCCTG 234  
 QY 181 ATCTCTCTGATGTGAGACCAAGAACCAAGTCAAGAGCAATGAGAACTCTCT 240  
 DB 235 ATCTCTCTGATGTGAGACCAAGAACCAAGTCAAGAGCAATGAGAACTCTCTCT 294  
 QY 241 CTGCACTGATCCGTCATCATTCATGATGCTGCTGATTTCTTAAGCTCCAAAGCGC 300  
 DB 295 CTGCACTGATCCGTCATCATTCATGATGCTGCTGATTTCTTAAGCTCCAAAGCGC 354  
 QY 301 TTGGTAACTGGCCCGGGGACTTTTGGCCGCTTCTGCGCTTGGGGGCAAAAGAGCC 360  
 DB 355 TTGGTAACTGGCCCGGGGACTTTTGGCCGCTTCTGCGCTTGGGGGCAAAAGAGCC 414  
 QY 361 GACTCTGCAAGTTCCAGAACCTGAGAAACCGGGCGAAAGCCGGGGCCCGGCGCG 420  
 DB 415 GACTCTGCAAGTTCCAGAACCTGAGAAACCGGGCGAAAGCCGGGGCCCGGCGCG 474  
 QY 421 GATTACTATCTCAAAACCTATGAGATGACATTTGTTGAGAGTGTGCTTAATAA 480  
 DB 475 GATTACTATCTCAAAACCTATGAGATGACATTTGTTGAGAGTGTGCTTAATAA 501  
 QY 481 CAAAGAGGCTGCAAGATTGGCCAGCACTGATGTTATGCTGTGCTCCAAAGAGAG 540  
 DB 502 ----- 501  
 QY 541 GAGAGCTCAGATGATCTGTTCTTCTAAGAGAGCAATCTCCAAATACCTAATCTTC 600  
 DB 502 ----- 501  
 QY 601 ATATTAAAGCCAGGCTCTAGACAACTTATTACTTGTGATGCTCTGAGCTCCAC 660  
 DB 502 -----GCCAGGCTCTAGACAACTTATTACTTGTGATGCTCTGAGCTCCAC 552  
 QY 661 CTGATATTGCTCTTAATGACATGCGCTGTAATCCCAATPACTCTGTAAGCAGTTCTAGT 720  
 DB 553 CTGATATTGCTCTTAATGACATGCGCTGTAATCCCAATPACTCTGTAAGCAGTTCTAGT 612  
 QY 721 GCCTTAGAGTGTGTAAGTACAGGCGCAGCAAAAGTAAACAATTTCTTGGGAAGCTGGGT 780  
 DB 613 GCCTTAGAGTGTGTAAGTACAGGCGCAGCAAAAGTAAACAATTTCTTGGGAAGCTGGGT 672

QY 781 AATGAGCCAAATTAATATGGAACCATGATGCGCGGGCAGTAAATGGCAGCCAGTTGGGA 840  
 DB 673 AATGAGCCAAATTAATATGGAACCATGATGCGCGGGCAGTAAATGGCAGCCAGTTGGGA 732  
 QY 841 AAGGATTAATCAATCCAGCTGAAGAGCCCTGTTGACAGCCCATCCGATTTATTCCAGAGCCAGC 900  
 DB 733 AAGGATTAATCAATCCAGCTGAAGAGCCCTGTTGACAGCCCATCCGATTTATTCCAGAGCCAGC 792  
 QY 901 TTAATAGCCCTTAATTTGGCGCGCGGAGGAAGAATGTATGCGCCCTCTAGATGATTC 960  
 DB 793 TTAATAGCCCTTAATTTGGCGCGCGGAGGAAGAATGTATGCGCCCTCTAGATGATTC 852  
 QY 961 ATGAAGGTGACAGAAAGTACATGATGACATGATGACATGATGACATGATGACATGATGAC 1020  
 DB 853 ATGAAGGTGACAGAAAGTACATGATGACATGATGACATGATGACATGATGACATGATGAC 912  
 QY 1021 CGGCTGCTCAAGGTGATGACATCTCTGAAAACCTGCGCTGTTAGACACATCTCTGACAC 1080  
 DB 913 CGGCTGCTCAAGGTGATGACATCTCTGAAAACCTGCGCTGTTAGACACATCTCTGACAC 972  
 QY 1081 ATTAGAAATTCAGAAAGGTGTTAATACATACATCTCAGAGAAAGAAATTTGCTTGA 1140  
 DB 973 ATTAGAAATTCAGAAAGGTGTTAATACATACATCTCAGAGAAAGAAATTTGCTTGA 1032  
 QY 1141 GGTGTGTGACCACTTCAAGTGAAGGACAAACAATCTATCCGATTCCTATGCTGACAGA 1200  
 DB 1033 GGTGTGTGACCACTTCAAGTGAAGGACAAACAATCTATCCGATTCCTATGCTGACAGA 1092  
 QY 1201 TTCTTATGTTGAACCTTTAGAAATGCTGGGCAATCAGGGCAATGATGCTGATACAG 1260  
 DB 1093 TTCTTATGTTGAACCTTTAGAAATGCTGGGCAATCAGGGCAATGATGCTGATACAG 1152  
 QY 1261 CACTCATTTTGTGACATGATACATCACTCTGTGACCAAGATTTTAAACCATTAACA 1320  
 DB 1153 CACTCATTTTGTGACATGATACATCACTCTGTGACCAAGATTTTAAACCATTAACA 1212  
 QY 1321 GACTACTGGCTCTCTCTCTCTTACAAAGGCTGATATGGCCCAAGCTTGGCTGGAT 1380  
 DB 1213 GACTACTGGCTCTCTCTCTCTTACAAAGGCTGATATGGCCCAAGCTTGGCTGGAT 1272  
 QY 1381 GTGGCTGGGCTCCAGCGGAAGCCAGCGCTGCGGAGTATCCGGGCAAACTAAGGAT 1440  
 DB 1273 GTGGCTGGGCTCCAGCGGAAGCCAGCGCTGCGGAGTATCCGGGCAAACTAAGGAT 1332  
 QY 1441 TATGCTACTGACAAACCAACCAACCACTAGCTTGTGGTGCATTAACCTTTT 1500  
 DB 1333 TATGCTACTGACAAACCAACCAACCACTAGCTTGTGGTGCATTAACCTTTT 1392  
 QY 1501 ATCATCACTTGCATGATCAAGAAAGAAATCAAGCTGCTGGAGCTCTCAGAGCAAG 1560  
 DB 1393 ATCATCACTTGCATGATCAAGAAAGAAATCAAGCTGCTGGAGCTCTCAGAGCAAG 1452  
 QY 1561 CTGGTTACCAAGTACTGCTGACAGCCCTATGGGCAAGAGGCTTAAAGTCCAAAGTCAAGT 1620  
 DB 1453 CTGGTTACCAAGTACTGCTGACAGCCCTATGGGCAAGAGGCTTAAAGTCCAAAGTCAAGT 1512  
 QY 1621 CAACTGAATGGCCAGCCCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
 DB 1513 CAACTGAATGGCCAGCCCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1572  
 QY 1681 CGCCCGCTTGGGCGCGGCGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
 DB 1573 CGCCCGCTTGGGCGCGGCGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1632  
 QY 1741 GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1779  
 DB 1633 GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1671

RESULT 8  
AF282886

AF282886

1922 bp

mRNA

linear

PRI 14-OCT-2000





Qy	1441	TATGCTCAGTCGCAAAACCAACCAACAACAACACTACGTCGTGCGATTCATTAACCTTTT	1500
Db	1327	TATGCTCACTGCGCAAAACCAACCAACAACAACAACAACACTACGTCGTGCGATTCATTAACCTTTT	1386
Qy	1501	ATCATCAACTTGCATTCGATCAAGAAAGAAATCAAGCTGCGTGGGACTCTTCAGAGACAG	1560
Db	1387	ATCATCAACTTGCATTCGATCAAGAAAGAAATCAAGCTGCGTGGGACTCTTCAGAGACAG	1446
Qy	1561	CTGGTTACACAGTACTGCTGCAACCCCTATGGGACGAGAGGCGCTTAAAGTCCAAATCACTAGT	1620
Db	1447	CTGGTTACACAGTACTGCTGCAACCCCTATGGGACGAGAGGCGCTTAAAGTCCAAATCACTAGT	1506
Qy	1621	CAACTGAAATGGCCAGCCCTTAGTGAATGGTGGACGACGGGAGACCTTCCAGAAATTGAAGCCC	1680
Db	1507	CAACTGAAATGGCCAGCCCTTAGTGAATGGTGGACGACGGGAGACCTTCCAGAAATTGAAGCCC	1566
Qy	1681	CGCCCCCTTCGGGCGCGCGGACATTTGGTCAATCCCTCCAGTCAACATGGGCGTTTATGTC	1740
Db	1567	CGCCCCCTTCGGGCGCGCGGACATTTGGTCAATCCCTCCAGTCAACATGGGCGTTTATGTC	1626
Qy	1741	GTCAGGAATGTCATAGCTTTGGCCCTGGCCGCAACCGATTA	1779
Db	1627	GTCAGGAATGTCATAGCTTTGGCCCTGGCCGCAACCGATTA	1665

RESULT	9		
AX286343			
LOCUS	AX286343	2326 bp	DNA
DEFINITION	Sequence 1 from Patent WO0181569.		linear
ACCESSION	AX286343		
VERSION	AX286343.1	GI:17048588	

REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	SOURCE
1	Heinrikson, R. L. and Bienkowski, M. J.	Heparatase II, a human heparatase paralog	Patent: WO 0181569-A 1 01-Nov-2001;	Pharmacia & Upjohn Company (US)	Location/Qualifiers
					1..2326
					/organism="Homo sapiens"
					/mol_type="unassigned DNA"
					/db_xref="taxon:9606"

	Query Match	Best Local Similarity	Matches 1603; Conservative	Score 1417.8	US 6	Length 2368
QY	1	ATGAGGGGAGCTTTGGAGCTTCCCTGAAAGCATGCCCTCAGAGCAATCCCGCCCCCGCG	90.1%;	0;	2;	Indels 174; Gaps 14
Db	25	ATGAGGGGAGCTTTGGAGCTTCCCTGAAAGCATGCCCTCAGAGCAATCCCGCCCCCGCG	90.1%;	0;	2;	Indels 174; Gaps 14
QY	61	TGCGTAGCCCCGGGGGGCTCTACTTTGGCTCTGTGGTCCATCTCCCTTCCACCGAG	90.1%;	0;	2;	Indels 174; Gaps 14
Db	85	TGCGTAGCCCCGGGGGGCTCTACTTTGGCTCTGTGGTCCATCTCTCCCTTCCACCGAG	90.1%;	0;	2;	Indels 174; Gaps 14
QY	121	GCTGAGAGCAGAGAGACCCCTGCTGCTGTAGACAGAGCTGCAGGTTTGAAGAAAAAGACCTG	90.1%;	0;	2;	Indels 174; Gaps 14
Db	145	GCTGAGAGCAGAGAGACCCCTGCTGCTGTAGACAGAGCTGCAGGTTTGAAGAAAAAGACCTG	90.1%;	0;	2;	Indels 174; Gaps 14
QY	161	ATTCTACTTAGTGTAGAGACCAAGAACCCAGTCAGAGACGTCAATGAGAACTTCCCTCT	90.1%;	0;	2;	Indels 174; Gaps 14
Db	205	ATTCTACTTAGTGTAGAGACCAAGAACCCAGTCAGAGACGTCAATGAGAACTTCCCTCT	90.1%;	0;	2;	Indels 174; Gaps 14
QY	241	CTGACAGTGGATCCGCTCCATCATTCATGTGTGCTGCATTTCTCTAAGCTCCAAAGCGC	90.1%;	0;	2;	Indels 174; Gaps 14
Db	265	CTGACAGTGGATCCGCTCCATCATTCATGTGTGCTGCATTTCTCTAAGCTCCAAAGCGC	90.1%;	0;	2;	Indels 174; Gaps 14

Qy	301	TTGGTAGCCTGGGCGGGAGACTTTGGCCGCTTTCTGAGCTTCGGGCGCAAAAGAAC	360
Db	325	TTGGTAGCCTGGGCGGGAGACTTTGGCCGCTTTCTGAGCTTCGGGCGCAAAAGAAC	384
Qy	361	GACTTCTGCAATTCAGAACCTGAGAACCCGGCGAAAAGCCGGGGGGCCGGGGCCG	420
Db	385	GACTTCTGCAATTCAGAACCTGAGAACCCGGCGAAAAGCCGGGGGGCCGGGGCCG	444
Qy	421	GATTACTATCTCAAAAACATATGAGATGACATTTTGGAGTGTGTTCCTTAGATAA	480
Db	445	GATTACTATCTCAAAAACATATGAGATGACATTTTGGAGTGTGTTCCTTAGATAA	504
Qy	481	CAGAAAGGCTGCAAGATTGGCCGACACCTGAGATTTAGTGTGGTGTCCAAAGGAGAG	540
Db	505	CAGAAAGGCTGCAAGATTGGCCGACACCTGAGATTTAGTGTGGTGTCCAAAGGAGAG	564
Qy	541	GCAGCTCAGATGCATCTGGTTCTTTCTAAGAGACAATTTCCAACTTACAGTAATCTC	600
Db	565	GCAGCTCAGATGCATCTGGTTCTTTCTAAGAGACAATTTCCAACTTACAGTAATCTC	624
Qy	601	ATATTACAGCAGAGCTCTAGACAAACTTAACTTGTCTGATTCCTCGACTCAAC	660
Db	625	ATATTACAGCAGAGCTCTAGACAAACTTAACTTGTCTGATTCCTCGACTCAAC	684
Qy	661	CTGATTTTGGCTTAATGCACTGCGTGTGTAATCCCAATAACTCTCGGAACAGTTCTAGT	720
Db	685	CTGATTTTGGCTTAATGCACTGCGTGTGTAATCCCAATAACTCTCGGAACAGTTCTAGT	744
Qy	721	GCCTGAGCTGTGTGAAGTACAGCGCAGCAAAAAGTACAACTTTCTTGGAACTGGGT	780
Db	745	GCCTGAGCTGTGTGAAGTACAGCGCAGCAAAAAGTACAACTTTCTTGGAACTGGGT	804
Qy	781	ATTATGCTTATGATCTATGCGATCGATGCTATGCGCGCGGCGATGATATGGCATGCGTGGG	840

Db	871	CGGGTGGTCAGAAAGTGATGGAACCTTCTCTGAAAACCTGGCTGTTCAGCAACACTCTCTGACGAG	930
Qy	1081	ATTAGAAAATTCAGAAAAGTGTTAATPACATACCTCCAGGAAAGAAATTTGGCTTGA	1140
Db	931	ATTAGAAAATTCAGAAAAGTGTTAATPACATACCTCCAGGAAAAGAAATTTGGCTTGA	990
Qy	1141	GATGATGAGCACCTCAGCTGAGGAGCAAAACATCTATCCGATTCCTATGCTGCAGGA	1200
Db	991	GGTGTGGTGAACCACTCAGCTGAGGAGCAAAACATCTATCCGATTCCTATGCTGCAGGA	1050
Qy	1201	TTCTTATGTTGAAACATTTAGAAATGCTGGCCATCAGGGCATTTGATGTCGTATACG	1260
Db	1051	TTCTTATGTTGAAACATTTAGAAATGTCGGCCATCAGGGCATTTGATGTCGTATACG	1110
Qy	1261	CACATCATTTTTCAGCAATGATATACATCACCTGTGTGAGCCAGAAATTTTAACCATTTACCA	1320
Db	1111	CACATCATTTTTCAGCAATGATATACATCACCTGTGTGAGCCAGAAATTTTAACCATTTACCA	1170
Qy	1331	GACATPCTGGCTCTCTCCCTCTACAAAGGCGCTGATCGGGCCCCCAAGTTTGGCTGTGAT	1380
Db	1171	GACATPCTGGCTCTCTCTCTCTACAAAGGCGCTGATCGGGCCCCCAAGTTTGGCTGTGAT	1230
Qy	1381	GTGGCTGGGCTCCAGCGGAAAGCCACGGCTGTGGCCGAGTATCCGGGACAAACTAAGGATT	1440





QY 1141 GGTGTGTGACACCTGAGCTGAGGACAAACATCTATCCATTCCTATGCTGACGA 1200  
 DB 1567 GGTGTGTGACACCTGAGCTGAGGACAAACATCTATCCATTCCTATGCTGACGA 1526  
 QY 1201 TTCTTATGTTGAACACTTTAGAAATGCTGAGCCATGAGGACATTTGATGATACGG 1260  
 DB 1527 TTCTTATGTTGAACACTTTAGAAATGCTGAGCCATGAGGACATTTGATGATACGG 1586  
 QY 1261 CACTCATTTTGTGACATGATATACATCACTCTGTGACACGAATTTTAACTTACCA 1320  
 DB 1587 CACTCATTTTGTGACATGATATACATCACTCTGTGACACGAATTTTAACTTACCA 1746  
 QY 1321 GACTACTGCT 1380  
 DB 1747 GACTACTGCT 1806  
 QY 1381 GTGCTGTGGCTTCCAGCGGAGCCACGCGCTTGGCGAGTATCCGGAACAACTTAAAGATT 1440  
 DB 1807 GTGCTGTGGCTTCCAGCGGAGCCACGCGCTTGGCGAGTATCCGGAACAACTTAAAGATT 1866  
 QY 1441 TATGCTCTGTCACAAACACCAACACCAACCACTAAGTGTGCTGCTGCTGCTGCT 1500  
 DB 1867 TATGCTCTGTCACAAACACCAACACCAACCACTAAGTGTGCTGCTGCTGCTGCT 1926  
 QY 1501 ATCATCACTTGCATTCATCAAGAAAGAAATCAAGCTGTGAGACTCTCAGACAG 1560  
 DB 1927 ATCATCACTTGCATTCATCAAGAAAGAAATCAAGCTGTGAGACTCTCAGACAG 1986  
 QY 1561 CTGTGTCCACGATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620  
 DB 1987 CTGTGTCCACGATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2046  
 QY 1621 CAACGTATGAGCCAGCCCTTATGATGATGATGATGATGATGATGATGATGATGATG 1680  
 DB 2047 CAACGTATGAGCCAGCCCTTATGATGATGATGATGATGATGATGATGATGATGATG 2106  
 QY 1681 CGCCCGCTTGGGGCGCGCGCGACATTTGATCATCTCTCAATCACTGAGCTTTTATG 1740  
 DB 2107 CGCCCGCTTGGGGCGCGCGCGACATTTGATCATCTCTCAATCACTGAGCTTTTATG 2166  
 QY 1741 GTCAAGATGTCATGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1799  
 DB 2167 GTCAAGATGTCATGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2205

RESULT 11  
 AK094037 2064 bp mRNA linear PRI 15-JUN-2002  
 LOCUS AK094037 Homo sapiens cDNA FLJ36718 f18, clone U78RU2010747, highly similar  
 DEFINITION to Homo sapiens heparanase-like protein HPA2b mRNA.  
 ACCESSION AK094037  
 VERSION AK094037.1 GI:21753015  
 KEYWORDS oligo cloning; f18 (full insert sequence).  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 AUTHORS  
 1 Suzuki, O., Sasaki, N., Aotsuka, S., Shoji, T., Ichihara, T.,  
 Shiohara, N., Matsunoto, K., Hirano, M., Sano, S., Nomura, R.,  
 Yoshikawa, Y., Matsumura, Y., Moriya, S., Chiba, E., Momiyama, H.,  
 Onogawa, S., Kaeriyama, S., Satoh, N., Matsumura, H., Takahashi, E.,  
 Katsukawa, R., Kuga, N., Kuroda, A., Satoh, I., Kamata, K., Takami, S.,  
 Terashima, Y., Watanabe, M., Sugiyama, T., Irie, R., Otsuki, T.,  
 Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,  
 Kanai, H., Nakamura, Y., Sekine, M., Kikuchi, H., Kanada, K.,  
 Maeshio, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanada, K.,  
 Wada, S., Murakawa, K., Kanehori, K., Takahashi, F., Fujii, A.,  
 Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,  
 Nagahara, K., Maeshio, Y., Nagai, K. and Isogai, T.  
 TITLE  
 JOURNAL  
 NEDO human cDNA sequencing project  
 Unpublished

REFERENCE 2 (bases 1 to 2064)  
 AUTHORS Isogai, T. and Yamamoto, J.  
 DIRECT SUBMISSION  
 JOURNAL Submitted (04-JUN-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7  
 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan  
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
 COMMENT NEDO human cDNA sequencing project supported by Ministry of  
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
 Research Association for Biotechnology (RAB); cDNA library  
 construction: Helix Research Institute (HRI) (supported by Japan  
 Key Technology Center etc.); 5' - & 3' - end one pass sequencing: RAB,  
 HRI, and Biotechnology Center, National Institute of Technology and  
 Evaluation; clone selection for full insert sequencing: HRI and  
 RAB; annotation: HRI and RAB.  
 FEATURES  
 source  
 1. 2064  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="U78RU2010747"  
 /issue\_type="clones"  
 /clone\_lib="UTERU2"  
 /note="Cloning vector: pME18FL3"  
 ORIGIN  
 Query Match 79.3%; Score 1411.4; DB 9; Length 2064;  
 Best Local Similarity 89.9%; Pred. No. 0;  
 Matches 1599; Conservative 0; Mismatches 6; Indels 174; Gaps 1;  
 QY 1 ATGAGGAGGAGCTTTGCTGCTTCCCTGAAGCCATGCGCCCTCAGACATCCGCGCCCGCG 60  
 DB 6 ATGAGGAGGAGCTTTGCTGCTTCCCTGAAGCCATGCGCCCTCAGACATCCGCGCCCGCG 65  
 QY 61 TGCCTAGCCCGGAGGAGCTTCTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
 DB 66 TGCCTAGCCCGGAGGAGCTTCTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 125  
 QY 121 GCTGAG 180  
 DB 126 GCTGAG 185  
 QY 181 ATTCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240  
 DB 186 ATTCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 245  
 QY 241 CTGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300  
 DB 246 CTGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 305  
 QY 301 TTGGTGAACCTGAGCCCGGAGAGCTTTCGCGCGCTTTCGCGCTTTCGCGCGAGAGAG 360  
 DB 306 TTGGTGAACCTGAGCCCGGAGAGCTTTCGCGCGCTTTCGCGCTTTCGCGCGAGAGAG 365  
 QY 361 GACTTCTGCAATTCAG 420  
 DB 366 GACTTCTGCAATTCAG 425  
 QY 421 GATTACTATCTCAAAACATGATGATGATGATGATGATGATGATGATGATGATGATG 480  
 DB 426 GATTACTATCTCAAAACATGATGATGATGATGATGATGATGATGATGATGATGATG 485  
 QY 481 CAGAAAGGCTGCAAG 540  
 DB 486 CAGAAAGGCTGCAAG 545  
 QY 541 GAGAGCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
 DB 546 GAGAGCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 605  
 QY 601 ATTCTAAGCAGAGAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
 DB 606 ATTCTAAGCAGAGAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 614

QY 661 CTGATATTGCTTAATGCACTGCTGTGTAATCCCAATAACTCTGGAAAGTTAGT 720  
 Db 615 ----- 614  
 QY 721 GCCCTGAGTCTGTGTAAGTACAGCGCCAGCAAAAAGTACAATTCTTGGGAACTGGGT 760  
 Db 615 ----- 614  
 QY 781 AATGAGCCAAATTAATCTATGCAACCATGATGGCGGGAGTAATGAGAGCAGTTGGGA 840  
 Db 615 ---GAGCCAAATTAATCTATGCAACCATGATGGCGGGAGTAATGAGAGCAGTTGGGA 671  
 QY 841 AAGATTATCATCAGCTGAAGAAGCTGTGGCAAGCCCATCGGATTATTCAGAGCCAGC 900  
 Db 672 AAGGATTATCATCAGCTGAAGAAGCTGTGGCAAGCCCATCGGATTATTCAGAGCCAGC 731  
 QY 901 TTATATGAGCCCTTAATATTGGGCGGGCCAGAGAAATGTCATCGCCCTCTAGATGATTC 960  
 Db 732 TTATATGAGCCCTTAATATTGGGCGGGCCAGAGAAATGTCATCGCCCTCTAGATGATTC 791  
 QY 961 ATGAAAGGTGAGAGAGTAAGTATGATGATGATGATGATGATGATGATGATGATGATG 1020  
 Db 792 ATGAAAGGTGAGAGAGTAAGTATGATGATGATGATGATGATGATGATGATGATGATG 851  
 QY 1021 CGGGGTGTCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080  
 Db 852 CGGGGTGTCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 911  
 QY 1081 ATTAGGAAATTCAGAAAGGTGATTAATCATACACTCCAGGAAAGAAATTTGGCTTGA 1140  
 Db 912 ATTAGGAAATTCAGAAAGGTGATTAATCATACACTCCAGGAAAGAAATTTGGCTTGA 971  
 QY 1141 GGTGTGTGACACCTCTGCTGAGGAGCAAAATATCTATCCATTTGCTGAGGGA 1200  
 Db 972 GGTGTGTGACACCTCTGCTGAGGAGCAAAATATCTATCCATTTGCTGAGGGA 1031  
 QY 1201 TTCTATGTTGAAACACTTTAGATGATGATGATGATGATGATGATGATGATGATGATG 1260  
 Db 1032 TTCTATGTTGAAACACTTTAGATGATGATGATGATGATGATGATGATGATGATGATG 1091  
 QY 1261 CACTCATTTTGGACATGATTAATCACTCTGTCGACCAAGATTTTAACTTAACTTAACT 1320  
 Db 1092 CACTCATTTTGGACATGATTAATCACTCTGTCGACCAAGATTTTAACTTAACTTAACT 1151  
 QY 1321 GACTACTGCT 1380  
 Db 1152 GACTACTGCT 1211  
 QY 1381 GTGGTGGGCTTCAGCGGAGCCAGCGCTGCGAGTATCCGGGCAAACTTAAAGATT 1440  
 Db 1212 GTGGTGGGCTTCAGCGGAGCCAGCGCTGCGAGTATCCGGGCAAACTTAAAGATT 1271  
 QY 1441 TATGTCATGTCAGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1500  
 Db 1272 TATGTCATGTCAGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1331  
 QY 1501 ATCATCAACTTGCATGATCAAGAAAGAAATCAAGCTGGCTGGGACTCTCAAGACAAG 1560  
 Db 1332 ATCATCAACTTGCATGATCAAGAAAGAAATCAAGCTGGCTGGGACTCTCAAGACAAG 1391  
 QY 1561 CTGGTTCACAGTACCTGCTGACGCTTATGGGAGAGAGGAGGAGGAGGAGGAGGAGGAG 1620  
 Db 1392 CTGGTTCACAGTACCTGCTGACGCTTATGGGAGAGAGGAGGAGGAGGAGGAGGAGGAG 1451  
 QY 1621 CAACTGAAATGAGCAGCCCTTATGATGATGATGATGATGATGATGATGATGATGATG 1680  
 Db 1452 CAACTGAAATGAGCAGCCCTTATGATGATGATGATGATGATGATGATGATGATGATG 1511  
 QY 1681 CGCCCCCTTGGGGCGGGCGGAGCATTTGTCATCCCTCAAGTCAACCAAGGGCTTTTATG 1740  
 Db 1512 CGCCCCCTTGGGGCGGGCGGAGCATTTGTCATCCCTCAAGTCAACCAAGGGCTTTTATG 1571  
 QY 1741 GTCAAGATATGCAATGCTTTGGCTGCGCTTACCGATTA 1779

Db 1572 GTCAAGATATGCAATGCTTTGGCTGCGCTTACCGATTA 1610  
 RESULT 12  
 AX393922  
 LOCUS 1673 bp DNA linear PAT 23-MAR-2002  
 DEFINITION Sequence 3 from Patent WO0204645.  
 AX393922  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 1 David, G. and Duerst, J.  
 A second human heparanase, and splice variants thereof, with a  
 predominant expression in skeletal muscle, heart and pancreas  
 Patent: WO 0204645-A 3 17-JAN-2002;  
 Vlaams Interuniversitair Instituut voor Biotechnologie vzw. (BI)  
 JOURNAL  
 FEATURES  
 source  
 CDS  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 85..1659  
 /note="unassigned protein product"  
 /codon\_start=1  
 /protein\_id="CAD29060.1"  
 /db\_xref="GI:19701883"  
 /db\_xref="EMBL:CAD29060"  
 /translation="MPSNSRPRPACLPAGATYLLALHLSLSQAGRRPLPYRAAG  
 LKEKTLILDVSTKSPRTVNEFLSLQDYSIIHDEWLDLSKRLVTLARGSPAR  
 LRFGRKADFLQFONLRNPAKSRGGPDYLYKNVEDIYASDVALXKQCKIARIP  
 DVMLELQREKAQMHVLYLKEQFNTSNLSLITPEPNVYRTHGVAVNGSQGXDYIDL  
 KSLQPIRISYGRASLYGENIGRPKNVYALDGMKVASVVDVAVTVOHGYIDRNVK  
 VMDELKTRLDLTSLDQIRKIQVNVYTPGKKIMLEGVYTSAGTNNLSQVAGFL  
 WNLNLTGMANGNVIVTRHSPFDGVMHNDONNPLPDVWLSILYRLLIKRYLAVH  
 VAGLQKRPGRVIRVDKIRITARYRPHCTNNHNYVRSITLFLINHSRKKIKLAGTIR  
 DKLVHOYLLQPYGEGEKSQVQLNGDPLVWVDGTLPELNPRLRGRITLVIPVTM  
 GEFVKNNALACRYR"  
 ORIGIN  
 Query Match 79.3%; Score 1411; DB 6; Length 1673;  
 Best Local Similarity 89.8%; Pred. No. 0;  
 Matches 1598; Conservative 1; Mismatches 6; Indels 174; Gaps 1;  
 QY 1 ATGAGGGTGTCTTTGTGCTTCTCTGAAGCCATGCTCCAGCAACTCCGCCGCCGCG 60  
 Db 55 ATGAGGGTGTCTTTGTGCTTCTCTGAAGCCATGCTCCAGCAACTCCGCCGCCGCG 114  
 QY 61 TGGCCTAGCCCGGGGGGCTCTACTTGGCTGTGTGCTCCATCTCTCCCTTCTCCAG 120  
 Db 115 TGGCCTAGCCCGGGGGGCTCTACTTGGCTGTGTGCTCCATCTCTCCCTTCTCCAG 174  
 QY 121 GCTGAGAGCAGAGAGACCTTCTCTGTAGACAGAGCTGAGGTTGAGAGAAAGACCTG 180  
 Db 175 GCTGAGAGCAGAGAGACCTTCTCTGTAGACAGAGCTGAGGTTGAGAGAAAGACCTG 234  
 QY 181 ATTCTACTGATGTGAGAGACCAAGAACCCATGAGAGAGTCAATGAGAACTTCTCTCT 240  
 Db 235 ATTCTACTGATGTGAGAGACCAAGAACCCATGAGAGAGTCAATGAGAACTTCTCTCT 294  
 QY 241 CTGAGCTGATTCGCTCATCTATGATGATGATGATGATGATGATGATGATGATGATG 300  
 Db 295 CTGAGCTGATTCGCTCATCTATGATGATGATGATGATGATGATGATGATGATGATG 354  
 QY 301 TTGGTGAACCTTGGGGCGGGAGCTTTCGCGGCTTTCGCGCTTCGCGGGGAAAGAGACC 360  
 Db 355 TTGGTGAACCTTGGGGCGGGAGCTTTCGCGGCTTTCGCGCTTCGCGGGGAAAGAGACC 414  
 QY 361 GACTTCTGCAATTCAGAACTTGAAGAAACCCGCGGAAAGACGCGGGGGGCCCGGCGG 420

Db	415	GACTTCCTGCAAGTTCCAGAACTCTGAGGACC	CGCGGAAAAGCCGCGGGGGCCCGGGCCCG	474
Oy	421	GATTACTATCTCCAAAACCTATGAGATGACAT	TTGTGTGAAGTGA TGTGTGCTTAGATATAA	480
Db	475	GATTACTATCTCCAAAACCTATGAGATGACAT	TTGTGTGAAGTGA TGTGTGCTTAGATATAA	534
Oy	481	CAGAAAAGGCTGCAAGATTGGCCGACACCC	CTGATGTATGTGCTGTGAGCTCCAAAGGAG	540
Db	535	CAGAAAAGGCTGCAAGATTGGCCGACACCC	CTGATGTATGTGCTGTGAGCTCCAAAGGAG	594
Oy	541	GCAGCTCAGATGCACTGTGTTCTTTCTAA	AGGAGCAATTCCTCAATACTTTCACGTAATCTC	600
Db	595	GCAGCTCAGATGCACTGTGTTCTTTCTAA	AGGAGCAATTCCTCAATACTTTCACGTAATCTC	654
Oy	601	ATATTTAACAGCCAGAGTCTCTAGACAAAC	TTTATTAATCTTTCAGTTCGTCGTGACCTCAC	660
Db	655	ATATTTAAC		663
Oy	661	CTGATATTTGTCTTAATATGCACTGCGTC	GTATAATCCCAATAACTCCTTGAAACAGTTCTAGT	720
Db	664			663
Oy	721	GCCTTGAGTCTGTGAAAGTACAGCGCAG	CAGAAAAGTACACATTTCTTGGGAACTGGGT	780
Db	664			663
Oy	781	AATAGCCAAATAACTATCCGACCATATGA	TGGCCGGGAGTAATATGGCAGCCAGTTGGGA	840
Db	664	--GAGCCAAATTAATCTATCCGACCATATGA	TGGCCGGGAGTAATATGGCAGCCAGTTGGGA	720
Oy	841	AAGGATTAACATCCAGCTGAAAGAGCC	CTGTGGACCCATCCGGAATTTATCCAGAGCCAGC	900
Db	721	AAGGATTAACATCCAGCTGAAAGAGCC	CTGTGGACCCATCCGGAATTTATCCAGAGCCAGC	780
Oy	901	TTATATGGCCCTTAATATTTGGGCGCG	CCGAGGAAAGATGTCATCGCCCTCTAGATGGAATTC	960
Db	781	TTATATGGCCCTTAATATTTGGGCGCG	CCGAGGAAAGATGTCATCGCCCTCTAGATGGAATTC	840
Oy	961	ATGAAGGTGGCAGGAAAGTACAGTAGATG	ATACCTTCGACCAACATTTGCTACATTTGATGTC	1020
Db	841	ATGAAGGTGGCAGGAAAGTACAGTAGATG	ATACCTTCGACCAACATTTGCTACATTTGATGTC	900
Oy	1021	CGGGTGTGCAAGGTGATGAGCACTTCCT	GTAAACTCGGCCCTGTAGACACACTCTCGACGAG	1080
Db	901	CGGGTGTGCAAGGTGATGAGCACTTCCT	GTAAACTCGGCCCTGTGTAGACACACTCTCGACGAG	960
Oy	1081	ATTAGGAATAATCCAGAAAGGTATTAAT	CATATACTCCAGAGAAAGAAATTTGGCTTGA	1140
Db	961	ATTAGGAATAATCCAGAAAGGTATTAAT	CATATACTCCAGAGAAAGAAATTTGGCTTGA	1020
Oy	1141	GGTGTGTGACCACTCCAGCTGAGGAGCA	CAAAACAATCTATCCGATTCCTATGCTGCAGGA	1200
Db	1021	GGTGTGTGACCACTCCAGCTGAGGAGCA	CAAAACAATCTATCCGATTCCTATGCTGCAGGA	1080
Oy	1201	TTCTTATGTGAAACACTTTAGAAATGCT	GTGCAATCAAGGCAATTTGATGTGTGATACGG	1260
Db	1081	TTCTTATGTGAAACACTTTAGAAATGCT	GTGCAATCAAGGCAATTTGATGTGTGATACGG	1140
Oy	1261	CACATCATTTTGAACATGATATACATAC	CTCGTGACAGAAATTTTAACCATTAACCA	1320
Db	1141	CACATCATTTTGAACATGATATACATAC	CTCGTGACAGAAATTTTAACCATTAACCA	1200
Oy	1321	GACTATCTGGCTCTCTCTCTCTATCAAG	CGGCTGATCGGCCCCAAAGTCTTGCTGTGCAT	1380
Db	1201	GACTATCTGGCTCTCTCTCTCTCTATCAAG	CGGCTGATCGGCCCCAAAGTCTTGCTGTGCAT	1260
Oy	1381	GTGGCTGGGGCTCCAGCGGAGCAACAGG	CTGGCCGAGATGCCGGA CAACCTAAGAGAT	1440
Db	1261	GTGGCTGGGGCTCCAGCGGAGCAACAGG	CTGGCCGAGATGCCGGA CAACCTAAGAGAT	1320
Oy	1441	TATGTCTACCTGCAAAACACACCAACA	CAACTAAGTTGTGTGGTCCATTACACTTTT	1500

[illegible]

QY 487 GGCTGAAGATGCGCCAGACCCCTGATGTTATGCTGCTGCTCCAAAGGAGAGAGAGCT 546  
 DB 361 GGCTGAAGATGCGCCAGACCCCTGATGTTATGCTGAGAGCTCCAAAGGAGAGAGAGCT 420  
 QY 547 CAGATGATCTGCTCTTCTTAAAGAGCAATCTCCAAATCTTAAGTAATCTCATATTA 606  
 DB 421 CAGATGATCTGCTCTTCTTAAAGAGCAATCTCCAAATCTTAAGTAATCTCATATTA 480  
 QY 607 ACAGCAGGCTCTAGACAAACTTATTAACCTTGCTGATGCTGAGACCTCCAGCTGATA 666  
 DB 481 ACA----- 483  
 QY 667 TTGCTCTAAATGCACTGCGCTGTAATCCAAATTAATCTCGAAGCACTTCTAGTCCCTG 726  
 DB 484 ----- 483  
 QY 727 AGTCTGTGAAGTACAGGCGCAGCAAAAGTACAACTTCTTGGAACCTGGGTAATGAG 786  
 DB 484 -----GAG 486  
 QY 787 CCAATTAATCTGCGACCATGATGCGCGGAGTAATG3CAGCCAGTGGAGAAAGAT 846  
 DB 487 CCAATTAATCTGCGACCATGATGCGCGGAGTAATG3CAGCCAGTGGAGAAAGAT 546  
 QY 847 TACATCCAGCTGAAGAGCTGTTGCGAGCCCATCCGGAATTTATTCAGAGCCAGCTTATAT 906  
 DB 547 TACATCCAGCTGAAGAGCTGTTGCGAGCCCATCCGGAATTTATTCAGAGCCAGCTTATAT 606  
 QY 907 GAGCCATATATGCGCGCGCGGAGAGAGATGTCACGCCCTCTAGATGAGTTCATGAAG 966  
 DB 607 GAGCCATATATGCGCGCGCGGAGAGAGATGTCACGCCCTCTAGATGAGTTCATGAAG 666  
 QY 967 GTGCGAGAGTACAGTATGATGCACTTACCTGCGCAACATTCATGATGATGAGTGGCCGGGTG 1026  
 DB 667 GTGCGAGAGTACAGTATGATGCACTTACCTGCGCAACATTCATGATGATGAGTGGCCGGGTG 726  
 QY 1027 GTCAAGGTGATGAGCTTCTGTAAGAACTGCGCTGTTAGACAACCTCTGACAGATTAAG 1086  
 DB 727 GTCAAGGTGATGAGCTTCTGTAAGAACTGCGCTGTTAGACAACCTCTGACAGATTAAG 786  
 QY 1087 AAAATTCGAAAGTGTGTATATACATACCTTCAGAGAAAGAGATTTGGCTTGAAGGTGTG 1146  
 DB 787 AAAATTCGAAAGTGTGTATATACATACCTTCAGAGAAAGAGATTTGGCTTGAAGGTGTG 846  
 QY 1147 GTGACCACTGAGCTGAGAGGAGCAAAACAATCTATCCGATTCCTATGCTGAGAGATCTTA 1206  
 DB 847 GTGACCACTGAGCTGAGAGGAGCAAAACAATCTATCCGATTCCTATGCTGAGAGATCTTA 906  
 QY 1207 TGGTTGAACACTTTAGAAATGCTGGCAATCGAGGCAATTAATGCTGATACGCACTCA 1266  
 DB 907 TGGTTGAACACTTTAGAAATGCTGGCAATCGAGGCAATTAATGCTGATACGCACTCA 966  
 QY 1267 TTTTGTGACATGATACATACCTGCTGAGACCAAGATTTTAAACCATTAACAGACTAC 1326  
 DB 967 TTTTGTGACATGATACATACCTGCTGAGACCAAGATTTTAAACCATTAACAGACTAC 1026  
 QY 1327 TGGCTCTCTCTCTCAAGAGGCTGATCGGCCCAAGGCTTGGCTGTGCAATGAGCT 1386  
 DB 1027 TGGCTCTCTCTCTCAAGAGGCTGATCGGCCCAAGGCTTGGCTGTGCAATGAGCT 1086  
 QY 1387 GGGCTCCAGCGGAGCCAGCGGCTGCGGAGTATCCGGGACAAACTTAAGATTTATGCT 1446  
 DB 1087 GGGCTCCAGCGGAGCCAGCGGCTGCGGAGTATCCGGGACAAACTTAAGATTTATGCT 1146  
 QY 1447 CACTGCAAAACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1506  
 DB 1147 CACTGCAAAACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1206  
 QY 1507 AACTTGATGATCAAGAAAGAAATCAAGCTGGCTGGAGCTCTCAAGACCAAGCTGGT 1566  
 DB 1207 AACTTGATGATCAAGAAAGAAATCAAGCTGGCTGGAGCTCTCAAGACCAAGCTGGT 1266

QY 1567 CACAGTACTGCTGCTGACGCCCTATGAGGAGGAGGCTAAAGTCCAACTGAGTGAATG 1626  
 DB 1267 CACAGTACTGCTGCTGACGCCCTATGAGGAGGAGGCTAAAGTCCAACTGAGTGAATG 1326  
 QY 1627 AATGCGCAGCCCTTATGATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1686  
 DB 1327 AATGCGCAGCCCTTATGATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1386  
 QY 1687 CTTGCGGCGCGGCGAGCACTTGGTCAATCCCTCAGTACACATAGGAGGCTTTATGAGT 1746  
 DB 1387 CTTGCGGCGCGGCGAGCACTTGGTCAATCCCTCAGTACACATAGGAGGCTTTATGAGT 1446  
 QY 1747 AATGCAATGCTTTGGCTGCGCTACCGATTA 1779  
 DB 1447 AATGCAATGCTTTGGCTGCGCTACCGATTA 1479  
 RESULT 14  
 AX280370  
 LOCUS AX280370 2369 bp DNA linear PAT 02-NOV-2001  
 DEFINITION Sequence 1 from Patent WO0177341.  
 ACCESSION AX280370  
 VERSION AX280370.1 GI:16607748  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1  
 Smets, G.M. and Sprengel, J.J.  
 Mammalian heparinase  
 TITLE Patent: NO 0177341-A 1.18-OCT-2001;  
 JANSSEN PHARMACEUTICA N.V. (BE)  
 FEATURES  
 source  
 1..2369  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Query Match 63.6%; Score 1132.2; DB 6; Length 2369;  
 Best Local Similarity 87.9%; Pred. No. 4.2e-306;  
 Matches 1321; Conservative 0; Mismatches 8; Indels 174; Gaps 1;  
 QY 277 CTGATTTCCCAAGTCCAGAGGCTTGTGAGCCCTGAGCGGAGACTTGGCCGCCCTT 336  
 DB 108 CTCTTCCTCTGAGCTCCAGCGCTGTGTGACCTGCGCGGAGACTTGGCCGCCCTT 167  
 QY 337 CTGCGCTTGGGCGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 396  
 DB 168 CTGCGCTTGGGCGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 227  
 QY 397 AAAAGCGCGGCGGCG 456  
 DB 228 AAAAGCGCGGCGGCG 287  
 QY 457 CGAAGTATGTTGCTTATGATTAACAGAAAGCTGCAAGATTGGCCAGACCTGATGTT 516  
 DB 288 CGAAGTATGTTGCTTATGATTAACAGAAAGCTGCAAGATTGGCCAGACCTGATGTT 347  
 QY 517 ATGCTGAGCTCCAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 576  
 DB 348 ATGCTGAGCTCCAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 407  
 QY 577 TTCTCAATACTTACAGTATCTCATATTAACAGCAGAGCTCTAGACAAACTTTATTAAC 636  
 DB 408 TTCTCAATACTTACAGTATCTCATATTAAC----- 440  
 QY 637 TTGCTGATGCTCTGAGACTCAACTGATATTTGCTCTAAATGACAGCTGCTGATATCC 696  
 DB 441 ----- 440  
 QY 697 AATACTCTGGAACAGTTCTAGTCCCTGAGTCTGTGAAGTACAGCGCCAGCAAAAG 756

Db	441	-----	440
Oy	757	TACAAcATTCTTGGGAaCTGGGTATaTGAAGCAATTAaCTATCCGAaCAATGCATAGCCGG	816
Db	441	-----	473
Oy	817	GCAGTAATAGCAGcCAGTGTGGAAaAGATTACATCCAGCTGAAGAGCCTGTGCAGCC	876
Db	474	CGAGTAATAGCAGcCAGTGTGGAAaAGATTACATCCAGCTGAAGAGCCTGTGCAGCC	533
Oy	877	ATCCGANTTATTCAGAGcCAGCTTATATGGCCCTAATTTGGCGCCCGAGGAAGAT	936
Db	534	ATCCGANTTATTCAGAGcCAGCTTATATGGCCCTAATTTGGCGCCCGAGGAAGAT	593
Oy	937	GTATAGCCCTCTAGATGSAaTTCAATGAAGGTGGCAGGAAGATACAGTATGATGCTTACC	996
Db	594	GTATAGCCCTCTAGATGSAaTTCAATGAAGGTGGCAGGAAGATACAGTATGATGCTTACC	653
Oy	997	TGGCAACAaTTGCTATCATTTAGTATGCGCGGGTGGTCAAGGTATGGAaCTTCTGAAAaCTGGC	1056
Db	654	TGGCAACAaTTGCTATCATTTAGTATGCGCGGGTGGTCAAGGTATGGAaCTTCTGAAAaCTGGC	713
Oy	1057	CTGTAGAaCAaCTCTCGACCAAGATTAAGAAAaTTCAAGAGTGGTAAATCAATACaCT	1116
Db	714	CTGTAGAaCAaCTCTCGACCAAGATTAAGAAAaTTCAAGAGTGGTAAATCAATACaCT	773
Oy	1117	CCAGAAAAGAAaTTTGGCTTGAAGGTGTGTGAaCCACCTCAaGCTGAAGCAAAaCAAT	1176
Db	774	CCAGAAAAGAAaTTTGGCTTGAAGGTGTGTGAaCCACCTCAaGCTGAAGCAAAaCAAT	833
Oy	1177	CTATCCGAaTTCTATAGCTGAGAGATTTCTATAGTGTGAaCACTTAAGAAaTGTGGCCAAaT	1236
Db	834	CTATCCGAaTTCTATAGCTGAGAGATTTCTATAGTGTGAaCACTTAAGAAaTGTGGCCAAaT	893
Oy	1237	CAGGcCATTTAGTGTGTATAGCGGCACTCAaTTTGTGAaCAATGAATACaCTCGTG	1296
Db	894	CAGGcCATTTAGTGTGTATAGCGGCACTCAaTTTGTGAaCAATGAATACaCTCGTG	953
Oy	1297	GACCGAAaTTTAAaCCCAaTTACCAaCTACaTGGCCTCTCTCCCTCAaAGCGCTGAaTc	1356
Db	954	GACCGAAaTTTAAaCCCAaTTACCAaCTACaTGGCCTCTCTCTCCCTCAaAGCGCTGAaTc	1013
Oy	1357	GcCCCCAAAGaGCTTGGCTGTGCATGTGGCTGAGCGGAaCCaCGGcCTTGGCCGA	1416
Db	1014	GcCCCCAAAGaGCTTGGCTGTGCATGTGGCTGAGCGGAaCCaCGGcCTTGGCCGA	1073
Oy	1417	GTTGATCCGGGAaCAaCTAAGATTTATGTCTCACTGCAaCAAAaCAaCAaCAaCAaCTAC	1476
Db	1074	GTTGATCCGGGAaCAaCTAAGATTTATGTCTCACTGCAaCAAAaCAaCAaCAaCAaCTAC	1133
Oy	1477	GTTGCTGGGTCCATTAaCACTTTTATCAaCTTGAaCTCAaCTGATCAaGAAGAAaATCAAG	1536
Db	1134	GTTGCTGGGTCCATTAaCACTTTTATCAaCTTGAaCTCAaCTGATCAaGAAGAAaATCAAG	1193
Oy	1537	CTGGCTGGGAaCTCTCAAGaCAaAGCTGGTTCAaCAaTCTGCTGCAaGCTTAATGGGCAaG	1596
Db	1194	CTGGCTGGGAaCTCTCAAGaCAaAGCTGGTTCAaCAaTCTGCTGCAaGCTTAATGGGCAaG	1253
Oy	1597	GAGGcCTTAAAGTCCAAGcCAGTGAaCTGAaCTGAATGcCAAGcCTTAATGATGTGTGAaCAaC	1656
Db	1254	GAGGcCTTAAAGTCCAAGcCAGTGAaCTGAaCTGAATGcCAAGcCTTAATGATGTGTGAaCAaC	1313
Oy	1657	GGAaCCCTCCCAaATTAAGcCCGcCCCTCTGGGcCGGcCGGAaCAaTTGTCATCCCT	1716
Db	1314	GGAaCCCTCCCAaATTAAGcCCGcCCCTCTGGGcCGGcCGGAaCAaTTGTCATCCCT	1373
Oy	1717	CGAGTCAaCATAGGcCTTTATGTGTCAaGAATGTCAaTGTCTTGGCTGCGCTCAaCGGA	1776
Db	1374	CGAGTCAaCATAGGcCTTTATGTGTCAaGAATGTCAaTGTCTTGGCTGCGCTCAaCGGA	1433
Oy	1777	TAA 1779	

DB	1434	TAA	1436
RESULT 15			
LOCUS	AF282885	1760 bp	mRNA
DEFINITION	Homo sapiens heparanase-like protein HPA2a	mRNA	complete cds.
ACCESSION	AF282885		
VERSION	AF282885.1	GI:10801194	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Humalovic, M., Nezarova, Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 1760)		
	McKenzie, E., Tyson, K., Stamps, A., Smith, P., Turner, P., Barry, R., Hitchcock, M., Patel, S., Barry, E., Stuberfield, C., Terrett, J. and Page, M.		
TITLE	Cloning and expression profiling of Hpa2, a novel mammalian heparanase family member		
JOURNAL	Biochem. Biophys. Res. Commun. 276 (3), 1170-1177 (2000)		
MEDLINE	20483645		
PUBMED	11027606		
REFERENCE	2 (bases 1 to 1760)		
AUTHORS	McKenzie, E.A., Tyson, K. and Stamps, A.		
TITLE	Direct Submission		
JOURNAL	Submitted (28-JUN-2000) Biology Group, Oxford Glycosciences, 10 The Quadrant, Abingdon Science Park, Abingdon OX14 3IS, England		
FEATURES			
SOURCE	Location/Qualifiers		
	1..1760		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/chromosome="10"		
	/map="10q23-q24"		
	61..1503		
	/note="splice variant"		
	/codon_start=1		
	/product="heparanase-like protein HPA2a"		
	/protein_id="AA033421.1"		
	/db_xref="GI:10801195"		
	/translation="MRYLCAPFPEAMPESNSRPPACIALPALYLLHLILSLSSQADPR RPLDPAAGLKEKTLILLDVSTAPPTVENNELSLDLPESLTHDCMLDLSKRLVLSKRLV TLAAGSPALFREGKRDTEFOPONTNLPKSRGSDGPKRQVAVLKYNEEPNNYRIMHGG AVNSOISGXOYLOKSLLOPRLIRASLIGPNRGRKXVALLDPMYAGSTVAV VTMHCYIDRVRVYMDPELKEKRLDLTSLDQIKRQVNTYTPKRTIMBEVYVTSAG GNNLSIDYAGFLMLNTLDELANQGIIDVYIRHSFPHGTHLVDQNPFLPDYVLSL LYKRLIPKVALAVNAVLQRPKRGRIIDRLRLYAHCTMHHNNVYRGSTLTININ HRSKRTIKLAGTLADKLVHQLYLLPQYBQGLKSKSVQLNDQPLVMVDGTLPELPPRR LRAGRTIVIPVTWGFVFNKVNVALRYE"		
ORIGIN			
Query Match	61.6%; Score 1095.4; DB 9; Length 1760;		
Best Local Similarity	81.1%; Pred. No. 8-8e-236;		
Matches 1442; Conservative	0; Mismatches 1; Indels 336; Gaps 1;		
QY	1 ATGAGGGTGTCTTGTGCTTTCCTCGTGAAGCCATGCGCTTCAGCAACCTCCGCGCGCGCG	60	
DB	61 ATGAGGGGTCTTGTGCTTTCCTCGTGAAGCCATGCGCTTCAGCAACCTCCGCGCGCGCG	120	
QY	61 TGCCATGAGCCCGGGGGGGCTCTACTTGGGCTCTGTGGCTCATCTCTCCCTTCCGCCAG	120	
DB	121 TGCCATGAGCCCGGGGGGGCTCTACTTGGGCTCTGTGGCTCATCTCTCCCTTCCGCCAG	180	
QY	121 GCTGAGACAGAGAACCTCTTGCTGCTGTAGACAGAGCTGAGAGTTTGAAGAAAGAACCTG	180	
DB	181 GCTGAGACAGAGAACCTCTTGCTGCTGTAGACAGAGCTGAGAGTTTGAAGAAAGAACCTG	240	
QY	181 ATTCTACTGTATGTAGAGACCAACCAAGTACAGAGACAGTCAATGAGAACTTCTCTCT	240	
DB	241 ATTCTACTGTATGTAGAGACCAACCAAGTACAGAGACAGTCAATGAGAACTTCTCTCT	300	
QY	241 CTGACGTGATCCGTCATATTCATGATGAGTGGCTGCATTTCTTAAGTCCAGAGCG	300	

```

Db      ||||| 301 CTGACGCTGGATCCCGTCACTCATATGCTGCGATTCCTTAAGCTCCAAACGCG 360
Qy      ||||| 301 TTGATGACCTGGCCCGGGGACCTTTGCGCCGCTTTCTGCGCTTCGCGGGGCAAAAGGAC 360
Db      ||||| 361 TTGGTGAACCTGGCCCGGGGACCTTTGCGCCGCTTTCTGCGCTTCGCGGGGCAAAAGGAC 420
Qy      ||||| 361 GACTTCTGCACTTCGACAACTGAGAACCCGCGGAAAGCCGCGGGGGCCCGGGCCG 420
Db      ||||| 421 GACTTCTGCACTTCGACAACTGAGAACCCGCGGAAAGCCGCGGGGGCCCGGGCCG 480
Qy      ||||| 421 GATTAATCTCAAAAACATGAGATGACATTTGTCAGAGTGAATGTTGCTTAGATAA 480
Db      ||||| 481 GATTAATCTCAAAAACATGAGG----- 505
Qy      ||||| 481 CAGAAAGGCTGCAGATTCGCCAGCACCCCTGATGTTATGCTGCTGCTCAAAAGGAGAG 540
Db      ||||| 506 ----- 505
Qy      ||||| 541 GCAGCTCAGATGATCTGCTTCTTAAGAGCAATTCGCAATTAAGTATCTC 600
Db      ||||| 506 ----- 505
Qy      ||||| 601 AATTTAAGCAGCGAGTCTCTAGAACAACTTTAATCTTTGCTGATGCTGAGCTCAC 660
Db      ||||| 506 ----- 505
Qy      ||||| 661 CTGATATTTGCTTAATGCACTGCGGTGTAATCCCAATACTCTCGAAGAGTTCTAGT 720
Db      ||||| 506 ----- 505
Qy      ||||| 721 GCCCTGAGTCTGTTGAGTACAGCGCCAGCAAAAAGTACAACATTTCTTGGAATCGGGT 780
Db      ||||| 506 ----- 505
Qy      ||||| 781 AATGAGCCAAATTAATCTATCGACCAATGCAATGCGCGGCAATTAATGCGCACTGATGGGA 840
Db      ||||| 506 -ATGAGCCAAATTAATCTATCGACCAATGCAATGCGCGGCAATTAATGCGCACTGATGGGA 564
Qy      ||||| 841 AAGATTAATCATCAGCTGAGAGAGCTGTTGCGAGCCCATCCGAAATTTATTCAGAGCCAGC 900
Db      ||||| 565 AAGATTAATCATCAGCTGAGAGAGCTGTTGCGAGCCCATCCGAAATTTATTCAGAGCCAGC 624
Qy      ||||| 901 TTATATGCGCTTAATATTTGGCGCGCGAGAGAAATGTCATGCGCTCTTAATGATGATTC 960
Db      ||||| 625 TTATATGCGCTTAATATTTGGCGCGCGAGAGAAATGTCATGCGCTCTTAATGATGATTC 684
Qy      ||||| 961 ATGAAGTGGCAGAGAGTACAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
Db      ||||| 685 ATGAAGTGGCAGAGAGTACAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 744
Qy      ||||| 1021 CGGCTGCTCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
Db      ||||| 745 CGGCTGCTCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 804
Qy      ||||| 1081 AATAGAAATTAATCAGAAAGTGGTTAATCATACATCCAGAAAGAAAGATTTGCTTGA 1140
Db      ||||| 805 AATAGAAATTAATCAGAAAGTGGTTAATCATACATCCAGAAAGAAAGATTTGCTTGA 864
Qy      ||||| 1141 GGTGTGTGACCACTGACCTGAGAGCAAAACAATCTATCTGATGCTGAGGA 1200
Db      ||||| 865 GGTGTGTGACCACTGACCTGAGAGCAAAACAATCTATCTGATGCTGAGGA 924
Qy      ||||| 1201 TTCTATGTTGAACAATTAGAAATGCTGGCAATTCAGGGCAATTTGATGATGATGATGATG 1260
Db      ||||| 925 TTCTATGTTGAACAATTAGAAATGCTGGCAATTCAGGGCAATTTGATGATGATGATGATG 984
Qy      ||||| 1261 CACTCATTTTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
Db      ||||| 985 CACTCATTTTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1044
Qy      ||||| 1321 GACTACTGCTCTCTCTCTCAAGGCTGATGCGGCCCAAGTCTTGCTGAT 1380

```

```

Db      ||||| 1045 GACTACTGCTCTCTCTCTCAAGGCTGATGCGGCCCAAGTCTTGCTGAT 1104
Qy      ||||| 1381 GTGGCTGGCTCCAGCGAGGCAAGCCAGGCTGCGCAATGATCCGGGACAACTAAGAT 1440
Db      ||||| 1105 GTGGCTGGCTCCAGCGAGGCAAGCCAGGCTGCGCAATGATCCGGGACAACTAAGAT 1164
Qy      ||||| 1441 TATGCTCACTGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1500
Db      ||||| 1165 TATGCTCACTGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1224
Qy      ||||| 1501 ATCATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
Db      ||||| 1225 ATCATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1284
Qy      ||||| 1561 CTGGTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
Db      ||||| 1285 CTGGTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1344
Qy      ||||| 1621 CAATGATGCGAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680
Db      ||||| 1345 CAATGATGCGAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1404
Qy      ||||| 1681 CGCCCTTTCGCGCGCGCGCGGACATTTGATGATGATGATGATGATGATGATGATGATGATG 1740
Db      ||||| 1405 CGCCCTTTCGCGCGCGCGCGGACATTTGATGATGATGATGATGATGATGATGATGATGATG 1464
Qy      ||||| 1741 GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1779
Db      ||||| 1465 GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1503

```

Search completed: May 13, 2004, 14:18:53  
 Job time : 6978 secs